

**Measuring Job Content:
Skills, Technology, and Management Practices**

Michael J. Handel
Department of Sociology and Anthropology
Northeastern University
E-mail: m.handel@neu.edu

September 2008

This work was supported by the National Science Foundation (grant number IIS-0326343), the Russell Sage Foundation, and the Wisconsin Alumni Research Fund. I thank Peter Cappelli, Harry Holzer, Arne Kalleberg, Joel Rogers, Nora Cate Schaeffer, Chris Winship, Erik Olin Wright, Jonathan Zeitlin, and participants in the Economic Sociology Seminar at the University of Wisconsin–Madison for their advice and comments on this project. I also thank Mary Ellen Colten, Carol Cosenza, and Anthony Roman of the Center for Survey Research at the University of Massachusetts Boston for their work in planning and fielding the survey of Skills, Technology, and Management Practices (STAMP).

IRP Publications (discussion papers, special reports, and the newsletter *Focus*) are available on the Internet. The IRP Web site can be accessed at the following address: <http://www.irp.wisc.edu>

Abstract

The conceptualization and measurement of key job characteristics has not changed greatly for most social scientists since the Dictionary of Occupational Titles and Quality of Employment surveys were created, despite their recognized limitations. However, debates over the roles of job skill requirements, technology, and new management practices in the growth of inequality have led to renewed interest in better data that directly addresses current research questions. This paper reviews the paradigms that frame current debates, introduces new measures of job content designed specifically to address the issues they raise from the survey of Skills, Technology, and Management Practices (STAMP), and presents evidence on validity and reliability of the new measures.

Measuring Job Content: Skills, Technology, and Management Practices

Over twenty-five years ago, research on the Dictionary of Occupational Titles (DOT) transformed the measurement of job characteristics, particularly those relating to skill (Spenner 1979; Cain and Treiman 1981). Criticism of the quality of DOT soon followed, but there has been little work within sociology to develop improved measures. To a lesser extent, the same could be said about an array of other job measures in the Quality of Employment Survey (QES), also conducted in the 1970s (Staines 1979).

Yet after a certain lull of interest resulting from this disappointment, skill requirements and other job characteristics have become the focus of attention again. The reasons for the current interest include the growth of earnings inequality, perceived changes in the nature of work, continuing racial and ethnic disadvantages in the labor market, persistent poverty, and doubts over the quality of the nation's education systems. These issues have engaged a range of sociologists, labor economists, education researchers, and policy analysts, among others. Claims that job skill requirements are rising ever faster because of the spread of technology and teams have become almost an article of faith for many educators, public officials, and the public (Johnston and Packer 1987; U.S. Department of Labor 1991).

However, one of the most notable aspects of these discussions is scarcity, thinness, or absence of data that speak directly to key issues. There is, in fact, little hard, representative data on what people actually do at work. Researchers have only a general sense of levels of job skill requirements and even less information on rates of change and the specific dimensions along which job skills are changing. There is no information on the level of complexity of workers' reading, writing, and math tasks in any data set, for instance. We do not know how many jobs require simple addition and subtraction skills and how many require calculus, whether there is any trend in specific job requirements, and, if so, the rates of change over time. Almost no American survey has equally strong coverage of job skill requirements, technology use, and employee involvement practices despite their presumed importance and interrelationships. Trend data are generally absent.

Indeed, the field still lacks a well-defined framework for measuring the job characteristics that are at the center of so much interest. Twenty-five years ago Kenneth Spenner (1983, p.825) complained that “Our conceptualization and measurement of skill are poor. Unidimensional, undefined concepts, nonmeasures, and indirect measures of skill have not served us well.” In some respects, the situation has remained largely unchanged (Borghans, Green, and Mayhew 2001; Goddard 2001).

Researchers agree that a better understanding of the substantive issues requires improved measurement. Because the substantive issues are manifold, they require a unified instrument. This paper introduces and evaluates an effort to solve these problems using the survey of Skills, Technology, and Management Practices (STAMP). STAMP is a nationally representative, two-wave, refreshed panel survey (n=2,304) that includes a largely new set of measures designed specifically to address current debates over the nature of work. Employed adults were selected randomly within households using standard random-digit dial telephone survey procedures and asked a detailed set of questions about their jobs between October 2004 and January 2006. The survey contains approximately 166 unique items related to various job characteristics and required 28 minutes on average to administer.

Section I of this paper reviews the multiple, interlocking debates that have converged on the need for better measures of job content and motivated the survey. Section II discusses a measurement strategy called explicit scaling that guided the effort to generate better measures. Section III presents evidence on the content, construct, and criterion validity and reliability of the measures, in the tradition of earlier work on the DOT (Section III) (Cain and Treiman 1981). In the process, the paper also offers a map for understanding the domains implicated in the current debates over work (see Table 1). Analyses using the survey to examine the particular substantive issues for which it was designed will appear in subsequent papers.

Table 1
STAMP Survey Content (N=number of items)

Basic Job and Organizational Information (N=12)
Occupation, industry, organizational position, organizational and job tenure, union membership, organizational size, organization type
Skill and Task Requirements (N=60)
Cognitive skills (N=48)
<i>Mathematics</i> (n=12)
<i>Reading</i> (n=8)
<i>Writing</i> (n=6)
<i>Forms and visual matter</i> (n=6)
<i>Problem-solving</i> (n=3)
<i>Education, experience, and training requirements</i> (n=9)
<i>Skill changes in previous three years</i> (n=4)
Interpersonal job tasks (n=8)
Physical job tasks (n=4)
Computer and Non-computer Technology (N=49)
<i>Computers</i> (n=26)
Frequency of use
Use of fourteen specific applications
Use of advanced program features, occupation-specific, and new software
Training times
Complexity of computer skills required
Adequacy of respondents' computer skills
Computer knowledge and experience in prior jobs among non-users
<i>Machinery and electronic equipment</i> (n=18)
Level of machine knowledge needed, training time
Set-up, maintenance, and repair
Automation, equipment and tool programming
<i>Other technology</i> (n=5)
Telephone, calculator, fax, bar code reader, and medical, scientific and lab equipment
<i>Technological displacement measures</i>
Employee Involvement Practices (N=18)
Job rotation and cross-training
Pay for skill
Formal quality control program
Teams activity levels, responsibilities, and decision making authority
Bonus and stock compensation
Autonomy, Supervision, and Authority (N=11)
Closeness of supervision, autonomy
Repetitive work
Supervisory responsibilities over others
Decision-making authority over organizational policies
Job Downgrading (N=15)
Downsizing and outsourcing
Reductions in pay and retirement and health benefits
Promotion opportunity, internal labor markets
Work load, pace, and stress
Strike activity
Job Satisfaction (N=1)

I. DEBATES OVER JOBS AND WORK: SUBSTANTIVE AND METHODOLOGICAL ISSUES

A. Post-Industrialism vs. Deskillling

The long-running debate between deskillling and post-industrial theories is well-known and need not be rehearsed here (Bell 1973; Braverman 1974; for reviews, see Attewell 1987; Form 1987). Here, it is sufficient to note that from a methodological perspective the debate eventually faced two major obstacles.

The frequent reliance on case studies limited generalizability. In the absence of a systemic perspective, studies might uncover skill changes in particular occupations but miss offsetting changes in related occupations not studied or the effects of changing relative shares of differently skilled occupations in the overall workforce. Case studies could be rich sources of information but were not well suited to understanding the net changes in skill requirements for the economy as a whole (Spenner 1983; Attewell 1989).

In addition, the lack of consistent methods and measures meant that one could not be certain that any individual study covered all important dimensions of skill, i.e., content validity was seldom addressed. More seriously, the lack of standard measures prevented cumulation of results across studies of different occupations and no convincing profile of the job structure or its changing shape over time were possible.

The response was greater use of standardized measures derived from the Dictionary of Occupational Titles (DOT) in conjunction with standard labor force surveys (Spenner 1979; Howell and Wolff 1991). The DOT was attractive because its ratings derived from in-person interviews and observations by trained job analysts. However, the DOT was devised by applied psychologists and practitioners for career counseling, not social science research. The jobs analyzed are a convenience sample and recent editions simply carry over most ratings from previous editions, often dating to the 1950s and 1960s, because re-rating jobs proved too costly. DOT ratings are also occupation means, which

wash out all within-occupation variation. In the absence of a strong sense of how to improve the situation, the debate more or less stalled by the late 1980s with calls for better data (Cain and Treiman 1981; Attewell 1990; Spenner 1990; Vallas 1990; U.S. Department of Labor 1993, p.20).

An enduring methodological contribution was the recognition that skill trends could be decomposed, at least conceptually, into between-occupation changes in employment shares and within-occupation changes in job task content. In practice, the former is captured easily with standard survey data, while the latter is largely unknown, which remains a significant barrier to understanding job skill levels and trends (Spenner 1983, pp.825f).

B. The Political Economy of Changing Labor Market Institutions

Quite separately, Barry Bluestone and Bennett Harrison's work initiated a line of research on job trends that remains central to current debates (Bluestone and Harrison 1982; Harrison and Bluestone 1988; Harrison 1994). They were the first to note the rise in earnings inequality since the late 1970s and attributed it to the decline of subordinate primary jobs and the growth of secondary labor market jobs. Economic crises in the 1970s and 1980s led management to regain profitability by lowering labor costs using a "lean and mean" strategy. This explanation focuses on the unraveling of institutional protections that employees gained in the postwar period, in contrast to subsequent human capital explanations described below. From a methodological standpoint, many key variables present few difficulties and are available in standard data sets (e.g., union density, contingent worker status, minimum wage levels, trade sensitivity).

However, other relevant variables remain scarce or poorly measured, such as vulnerability to downsizing and outsourcing, pay and benefit reductions (givebacks), declines in internal labor markets, overwork, and job stress (Graham 1993; Tilly and Tilly 1997; Green 2004, 2006). Although there have been scattered efforts recently to gather data on some of these questions, the last systematic effort was the third Quality of Employment Survey conducted in 1977, which continues to exert a strong influence on recent approaches to measuring concepts such as autonomy, work effort, and job stress.

C. Flexible Specialization's New Paradigm of High-Quality Jobs

Piore and Sabel (1984) proposed an alternative institutionalist account of the American economy's response to the crises of the 1970s and 1980s. Firms found they could not compete based on lower price and labor costs, and differentiated themselves qualitatively through customization, niche marketing, high quality, and continuous innovation. These strategies require firms to upgrade the skill content of lower-level jobs and adopt more participatory and democratic management practices, often modeled on Japanese quality control techniques and employee involvement (EI) practices long advocated by management reformers (e.g., McGregor 1957; Walton 1974).

The spread of computers reinforced skill upgrading by replacing manual with mental labor and reversing the process of job fragmentation. Computers also decentralized information and decision making into the hands of ordinary workers, reinforcing the value of EI practices (Hirschhorn 1984; Zuboff 1988).

Flexible specialization theory predicts associations between high skill demands, advanced technology, and self-directed teams, which are mutually reinforcing (Appelbaum and Batt 1994). By contrast, the lean and mean camp views computers as a source of increased worker surveillance and decreased autonomy, and consider employee involvement either a method of effort intensification or merely a token gesture (Graham 1993; Vallas 2003).

Again, the substantive details do not require further elaboration here (for reviews see Smith 1997, Vallas 1999). However, several methodological limitations have slowed further progress on these issues.

Most work on both sides of this controversy uses case study methods.¹ Nationally representative data with reliable measures of EI practices, related job characteristics, and employee characteristics remain scarce. Basic constructs, like the meaning of self-directed teams, are subject to vague or varying definitions. There is no real consensus regarding basic descriptive information, such as incidence rates for

¹Exceptions include Appelbaum et al. 2000, Osterman 2000, Capelli and Neumark 2001, Handel and Gittleman 2004.

EI practices, much less the causal relationships between skills, technology, EI, and other elements of the flexible specialization paradigm.

D. Human Capital Theory and Skill-Biased Technological Change

Following Bluestone and Harrison's work, neoclassical labor economists developed their own explanation of the growth in earnings inequality based on human capital theory. In this view, the rapid spread of computer technology since the early 1980s increased the demand for skilled labor, widening education differentials (Katz and Murphy 1992; Autor, Katz, and Krueger 1998). This theory of skill-biased technological change (SBTC) is the dominant explanation for inequality growth among labor economists and favored by some sociologists (e.g., Fernandez 2001), as well more popular writers, educators, and policy-oriented analysts.

The exact causal argument relating computers to skills and wages remains somewhat unsettled. Many researchers interpret the wage premium for computer use as evidence that computer hardware and software are complex, requiring a significant training investment and receiving wage returns (Krueger 1993; Borghans and ter Weel 2004; Dickerson and Green 2004; Dolton and Makepeace 2004). Others argue that computers may not be complex or difficult to learn in themselves, but may require more general cognitive skills for the reasons noted in the previous section (Levy and Murnane 1996; Autor, Levy, and Murnane 2002; Spitz-Oener 2006). By this route, mainstream labor economists have come to embrace the flexible specialization position on employee involvement practices within the firm (Siegel 1999; Caroli and Van Reenen 2001; Bresnahan, Brynjolfsson, and Hitt 2002; Shaw 2002). However, mainstream labor economics remains less sympathetic to institutional arguments from the political economy perspective that focus on declining worker bargaining power.

There is a large research literature on skill-biased technological change (for reviews see Handel 2003a, 2004). One notable criticism is that even though skill upgrading may be the historical trend, in order for computers to explain rising inequality, the rate of skill upgrading must have accelerated since

the 1970s in computer-intensive sectors, which points to the need for time series data (Mishel and Bernstein 1998).

However, in the absence of direct measures of job cognitive skill requirements, the debate over SBTC, like the earlier deskilling controversy, faces the prospect of diminishing returns. Most research uses very imperfect proxies for job skill demands, such as workers' own educational attainment, coarse occupational categories, or DOT scores. Valid, reliable, and easily interpreted direct measures of job skill requirements remain scarce despite their recognized centrality. Skill requirements are the focus of numerous claims and counter-claims, but rarely direct measurement. Even scarcer are consistent time series data necessary to test the acceleration hypothesis.²

Strategies to measure the skill demands of computer technology itself are limited mostly to dummy variables for computer use or the value of computer investment within industries, rather than more direct or specific measures of the complexity of computer skills used on the job.³

E. Additional Sources of Concern over Job Trends

Three additional areas of research underscore the level of interest in job characteristics and the need for improved measures.

1. *Education Research and Policy*

Many education researchers and policymakers believe that the quality and quantity of Americans' basic schooling is inadequate for the demands of the new economy (Bishop 1989; Barton and Kirsch 1990; Berryman 1993; Murnane and Levy 1996; Rosenbaum and Binder 1997).

Official reports take for granted that the demand for cognitive and teamwork skills is accelerating dramatically, reflecting popular renditions of theories discussed above (U.S. National Commission on

²For an exception that uses German data see Spitz-Oener (2006).

³For exceptions that use British data see Borghans and ter Weel (2004) and Dickerson and Green (2004).

Excellence in Education 1983; U.S. Department of Labor 1991; for a review, including contrary evidence, see Handel 2005b).

Policy reports usually conclude with a laundry list of academic goals considered essential to meeting future workplace needs, but not accompanied by any credible evidence on the proportions of jobs requiring different kinds of academic and other skills.

Similar debates over trends and possible mismatches in education and job skill requirements are occurring in Britain and, to a lesser extent, Canada and continental Europe (Keep and Mayhew 1996; Payne 1999; Krahn and Lowe 1998; McIntosh and Steedman 2000; Haahr et al. 2004).

2. *Poverty and Racial/Ethnic Inequality*

Poverty research has also drawn on theories of post-industrialism and SBTC. William Julius Wilson (1987, 1996) argued that employment problems of the inner-city poor are due partly to a mismatch between the relatively low education levels of job seekers and the higher skill demands of growing industries and occupations (see also Holzer 1996; Moss and Tilly 2001).

3. *Interpersonal Skills*

Finally, a diverse group of researchers has focused on the increasing importance of interpersonal skills at work.

Employment shifts away from blue-collar and manufacturing jobs and toward white-collar and service work increase the share of jobs requiring interpersonal interaction, rather than manual skills (Bell 1973; Reich 1991). The prominence of teamwork in the employee involvement literature adds to this concern.

Growing workplace interaction demands have been cited as a possible barrier to the employment of low-income minority workers with nonstandard communication repertoires, cultural capital, and personal styles (Wilson 1996; Moss and Tilly 2001; Holzer and Stoll 2001; cf. Swidler 1986).

Interpersonal demands are also central to the growing literature on emotional labor and the related concept of caring labor in the study of service work (Hochschild 1983; Steinberg 1990; Leidner 1993;

Wharton 1999; Steinberg and Figart 1999; Brotheridge and Lee 2003; Glomb, Kammeyer-Mueller, and Rotundo 2004; England 2005).

Finally, somewhat contrary to the preceding, there is a literature on highly routinized service jobs, such as fast food and call center work, that have very short-cycle and scripted interactions and in which the interpersonal demands are more ritualized. This literature represents the application of deskilling theory to service sector work (Leidner 1993; Buchanan 2003). In this case, the question is documenting the limited interactive skills involved in expanding service work, not the potential upgrading effect claimed by the other views. However, both the deskilling and emotional/caring labor perspectives treat interactive demands as a source of job stress distinct from the other kinds of overwork discussed by the lean and mean perspective.

Nevertheless, as with cognitive skills, there has been limited effort to measure reliably the level of interactive skills required across different jobs, despite wide interest (for exceptions see Steinberg and Figart 1999; Brotheridge and Lee 2003; Glomb et al. 2004).

F. Perceived Need for Effective Measures of Job Content

Social scientists across disciplines involved in these research areas agree there is a significant data gap. A prominent proponent of the SBTC hypothesis noted:

Our understanding of how computer-based technologies are affecting the labor market has been hampered by the lack of large representative data sets that provide good measures of workplace technology, worker technology use, firm organizational practices, and worker characteristics (Katz 2000, p.237).

A National Research Council committee came to a similar conclusion:

The evidence on information technology and inequality suggests that computers may be partly responsible for the relative increase in the demand for skilled, educated workers. However, thus far the measures of “skill” and “education” are fairly coarse.... Time series data, even short time series, would be especially valuable in clarifying the role of technology in some of the organizational changes that are observed (National Research Council 1998, pp.38,44).

The United States President's Information Technology Advisory Committee argued that, "The information revolution puts a premium on basic knowledge, not just information technology literacy, but basic skills in reading, writing, communication, and teamwork" (PITAC 1999, pp.12, 58ff). Concerned that disadvantaged groups may fall further behind, the report concluded:

We need more data and we need to understand the social, economic, and policy issues in much greater depth. The research that is required to develop this knowledge should be broad-based, long-term, and large-scale in its scope... (PITAC 1999, p.55).

The National Institute for Occupational Safety and Health (NIOSH) called for a new survey of work organization and job characteristics administered on a three- or five-year cycle that would give special attention to the improvement and standardization of measures.

Since the demise of the Quality of Employment Surveys of the 1960s and 1970s, there has been no way of determining how the demands of work may be changing, and how these demands vary from one industry, occupation, or population to another...Several strategies for improving surveillance of the organization of work can be proposed. The most desirable (and ambitious) approach would be to develop a stand-alone nationally representative survey of the organization of work. Such a survey might be modeled after and expand upon the former QES (Department of Health and Human Services. National Institute for Occupational Safety and Health 2002, pp. vi, 7f.).

Previous recommendations to conduct an updated QES to address current concerns were not implemented (Kalleberg 1986).

In contrast, Britain has at least two survey programs, the Skills Survey and Workplace Employee Relations Study, each administered in five waves since the 1980s, covering cognitive-skill demands, computer use, and employee involvement practices, among others (Felstead, Gallie, and Green 2002; Cully, Woodland, O'Reilly, and Dix 1999). Canada, Australia, Germany, and the European Union also have multi-year survey programs dealing with skills, workplace organization, and other job characteristics (Goddard 2001; Harley 2002; Statistics Canada 2004; Spitz-Oener 2006; Paoli 1992; Gallie 1997; Parent-Thirion et al. 2007).

The survey of Skills, Technology, and Management Practices (STAMP) was developed to address the need for a similar U.S. survey.

II. SUBSTANTIVE QUESTIONS AND METHODOLOGICAL CONSIDERATIONS

The preceding suggests that a new worker survey needs to address the following substantive questions:

1. How many and which jobs require what levels of various skills, computer competencies,⁴ and participation in employee involvement practices? In brief, what is the skill profile of the American job structure?
2. What are the functional and causal relationships between skill requirements, computer use, and employee involvement (EI)?
3. What are the effects of skills, computers, and EI on wages and other characteristics that define desirable or undesirable jobs (e.g., work intensity, promotions, layoffs, outsourcing, unionization, job satisfaction)?
4. What are the trends in skill requirements, technology, and employee involvement practices, their interrelationships, and their relationships to trends in the other outcomes mentioned above?

However, the considerable uncertainty over the conceptualization of key constructs like skill, and the dissatisfaction with previous approaches underscores the need to improve their definition and measurement. Part of the solution is a measurement approach that could be called explicit scaling.

Explicit scaling involves questions and response options that are objective, specific, correspond directly to researchers' objects of interest, and have absolute meanings for respondents. Questions are phrased in terms of facts, events, and behaviors, rather than attitudes, evaluations, and holistic judgments. Items are general enough to serve as common measures for the diverse range of jobs within the economy, but sufficiently concrete that they have stable meanings across respondents. Response options discriminate across a wide range of levels to avoid floor and ceiling effects, and use natural units when possible. Rating scales, vague quantifiers, and factor scores, which have arbitrary metrics and lack specific or objective referents, are a last resort.

Explicit scales are intrinsically desirable for researchers because they have definite meanings compared to measures that are more abstract or vague. One would expect them to reduce measurement

⁴For ease of exposition, "computer use" is sometimes used as a shorthand to refer to the broader category "computer and other technology use."

error because they limit the scope for respondents' subjective interpretations of the meaning of items and some of the self-enhancing biases in self-reports that might result.

A contrary example from the Quality of Employment Survey asked employees to indicate their level of agreement (strongly agree-strongly disagree) with the statement "My job requires a high level of skill" (Quinn and Staines 1979). Such measures are relatively common in the study of work within psychology and elsewhere (Karasek 1979; Cook et al. 1981, pp.170ff.; Kalleberg and Lincoln 1988; Glick, Jenkins, Gupta 1986; Bosma et al. 1997; Fields 2002, pp.72ff.).

This question leaves the concept of skill undefined and nonspecific (holistic). Respondents have to decide for themselves the meaning of 'skill' and the weights to give the different aspects of their job in making a summary judgment. The response categories provide no objective guidelines that respondents could use to match their particular job's characteristics with a response option. The QES question and response options are so general and abstract that researchers can never be sure what the numerical scores mean in any concrete sense, such as whether or not a given rating means a job requires a certain level of education or math skills, for example.

The form of the question also invites responses based on relative rather than absolute standards. Because most people are not familiar with the full range of occupations, they are likely to judge their own job's skill level in comparison to others that are close to their own, rather than using the full range of jobs in the economy as their frame of reference. The job analysis field has long wrestled with the problem of obtaining self-ratings based on an objective or common standard, which is necessary if the measures are to mean the same thing across people and jobs (Harvey 1991, p.83).

The DOT's use of trained job analysts avoided problems associated with self-reporting, but some of the most commonly used DOT scales, such as a job's relationship to data, people, and things, often do not correspond to obvious, unambiguous, or concrete concepts, and the different levels of the scales are not even clearly ordinal (see below) (U.S. Department of Labor 1991, pp.3-1; Harvey 1991, p.147). Despite the value labels, the scales have the appearance of using arbitrary metrics to distinguish different

levels. Consequently, their face validity and interpretability is lower than if one were designing a measurement scheme *de novo*.

**Levels of Complexity of Jobs' Relationship to Data, People, and Things
from *The Dictionary of Occupational Titles***

<i>Data</i>	<i>People</i>	<i>Things</i>
0 Synthesizing	0 Mentoring	0 Setting Up
1 Coordinating	1 Negotiating	1 Precision Working
2 Analyzing	2 Instructing	2 Operating-Controlling
3 Compiling	3 Supervising	3 Driving-Operating
4 Computing	4 Diverting	4 Manipulating
5 Copying	5 Persuading	5 Tending
6 Comparing	6 Speaking-Signaling	6 Feeding-Off Bearing
	7 Serving	7 Handling
	8 Taking Instructions-Helping	

Note: Codes with higher values indicate lower levels of job complexity.

The DOT has recently been replaced as a career counseling tool by the U.S. Department of Labor's Occupational Information Network (O*NET), which gathers data through standardized surveys of representative samples of employees and converts responses to occupational mean scores (Peterson et al. 1999, 2001; U.S. Department of Labor 2005).

Most O*NET questionnaire items also differ from explicit scaling principles in being abstract, vague, jargon-laden, and multi-barreled (see Figure 1). Even definitions intended to clarify the questions are often complex.⁵ Research in industrial/organizational (IO) psychology suggests that respondents have much greater difficulty answering these kinds of abstract, holistic rating questions than more concrete and specific items (Harvey 1991, pp.95ff.; Harvey 2004; Spector and Fox 2003).

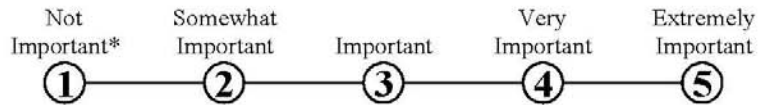
O*NET tries to avoid some of these problems with an anchored rating scale response format. Illustrative tasks are shown toward the lower, middle, and upper parts of the response continuum to create common reference points and help respondents select an answer. However, the illustrative tasks often seem difficult to apply to the general range of jobs. For example, it is hard to know how a police officer would decide whether their job required estimating skills on the level of determining the size of

⁵For the full text of O*NET surveys, see <http://www.onetcenter.org/questionnaires.html>.

FIGURE 1. SAMPLE O*NET QUESTIONNAIRE ITEMS

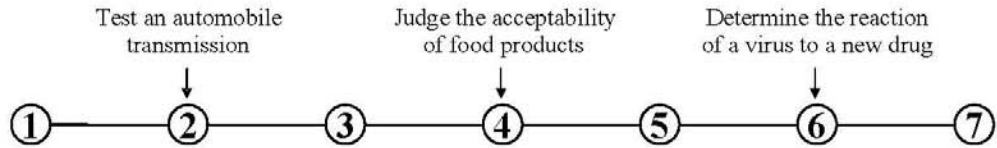
<p>2. Identifying Objects, Actions, and Events</p>	<p>Identifying information by categorizing, estimating, recognizing differences or similarities, and detecting changes in circumstances or events.</p>
---	---

A. How important is IDENTIFYING OBJECTS, ACTIONS, AND EVENTS to the performance of your current job?



* If you marked Not Important, skip LEVEL below and go on to the next activity.

B. What level of IDENTIFYING OBJECTS, ACTIONS, AND EVENTS is needed to perform your current job?



Highest Level

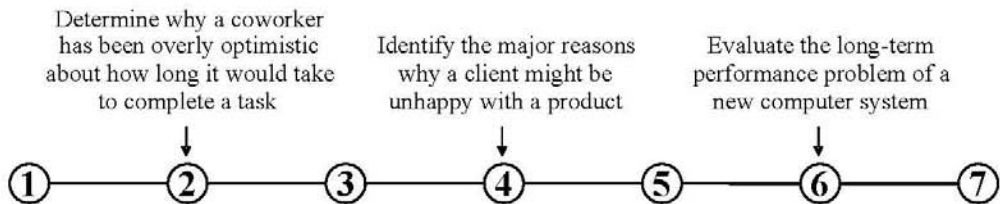
<p>30. Systems Evaluation</p>	<p>Identifying measures or indicators of system performance and the actions needed to improve or correct performance, relative to the goals of the system.</p>
--------------------------------------	---

A. How important is SYSTEMS EVALUATION to the performance of your current job?



* If you marked Not Important, skip LEVEL below and go on to the next skill.

B. What level of SYSTEMS EVALUATION is needed to perform your current job?



Highest Level

household furnishings to be crated (2), the time required to evacuate a city in the face of a major disaster (4), or the amount of natural resources beneath the world's oceans (6). The points on the scale do not seem to correspond to equal intervals, and the high complexity level of the task at the upper end creates the potential for scale compression and ceiling effects. The anchors do not really overcome the problem that rating scales have indefinite referents. Consequently, one can never be quite sure what O*NET scores actually mean in terms of specific real-world categories. If the level of 'systems evaluation' required by two occupations is 3 and 4, respectively, one cannot really explain how they differ on this dimension beyond the difference in scores themselves.

One of the benefits of explicit scaling is that it is possible to compare persons and jobs to determine if there is congruence or mismatch, which is not possible if job measures are on arbitrary scales without person measure equivalents. For example, one can easily compare job educational requirements with a person's own level of educational attainment because both are measured in a common, natural unit. By contrast, it is not easy to know whether workers are well-matched with jobs measured with most rating scales because they do not have clear, objective meanings and there are no corresponding person measures that also use them. This problem also affects the O*NET and DOT factor analytic scores that are often constructed from multiple rating scale items (Miller et al. 1980, pp.176ff.; Spenner 1990, p.403).

Realistically, creating measures that are behaviorally concrete and meaningful in absolute or objective terms is often difficult. Questions that are very precise, such as items on occupation-specific skills, may achieve clarity and concreteness at the cost of losing relevance for a large proportion of jobs. Some more general constructs resist explicit scaling because their manifestations are complex and qualitatively diverse across jobs, such as work intensity. As the table below suggests, even if a survey had space for a very long inventory of occupation-specific measures, there is no obvious way one could map the items to a common scale for full-sample analyses. The measures are incommensurate and any effort to equate them would be like comparing apples and oranges. Similar problems affect other commonly used concepts, such as job autonomy,

Occupation	Work Intensity Measure
Food service	Customers per hour
Assembly line worker	Parts worked per hour
Police officer	One- vs. two-person squad car
Teacher	Class size
Pilot	Continuous flying hours
Brand manager	Number of product lines
Physician	Patient load

task variety, and most forms of occupation-specific skill and training (e.g., operating printing machinery, writing a computer program in Perl). Avoiding rating scales in general labor force surveys can be difficult in such cases.

III. VALIDITY AND RELIABILITY OF STAMP MEASURES OF JOB CHARACTERISTICS

The application of explicit scaling concepts to the measurement of job characteristics in the STAMP survey can be evaluated by several standards of validity and reliability.

Explicit scales should exhibit particularly strong face validity and content validity. Their most visible characteristic is their clear meaning for both respondents and researchers. Given a clear definition of a construct (e.g., math use), content validity also means the measure(s) cover the different facets of the construct reasonably well, span the full range of the latent trait (i.e., no floor or ceiling effects), and exclude irrelevant constructs (Anastasi 1982). Content adequacy is based on the conceptual correspondence between measures and a construct's specified meaning. Readers can judge the face and content validity of STAMP items informally from the discussions below. Detailed comparisons between STAMP items and alternative measures from other surveys that support the choices made are available in a longer version of this paper (available from author).

Criterion validity refers to how well measures under consideration correlate with more direct or proven measures of the same construct (Anastasi 1982). In the absence of trained observer data on respondents' jobs, this paper uses the method of contrasted groups, which compares means across groups

for which they would be expected to differ (Anastasi 1982, pp.140f.; Bohrnstedt 1983, p.98). Criterion validity for the measures of job characteristics is high when they vary strongly in expected directions by occupation and education groups, and when validity coefficients are significantly higher between different measures and job educational requirements compared to personal educational attainment.

Construct validity refers to the extent to which a set of items measure a coherent, interpretively meaningful concept or trait (Anastasi 1982, pp.144ff.). One aspect of construct validity is the internal consistency or homogeneity of the items used to measure a construct. In this sense, Cronbach's α is a measure of construct validity as well as reliability, as are other measures of factor structure, such as principal components analysis and confirmatory factor analysis, all of which are used below (Schultz and Whitney 2005, p.120; Anastasi 1982, pp.146f.; Bohrnstedt 1983, pp.100ff.; Peterson et al. 1999).

A. Skills

Although the concept of skills has proven somewhat controversial, they are defined here as technical task demands that are defined by employers as necessary for effective job performance. STAMP divides the skill domain into cognitive, interpersonal, and physical skills, following the DOT's data, people, and things scheme. Cognitive skills are further divided into math, reading, writing, forms and visual matter, problem solving, and required education, job experience, and training times. Some of these correspond closely to school learning that has been the focus of great concern, while others relate to more general cognitive demands or are more job-specific.

1. *Math*

Mathematics is a relatively well-structured domain for which it was relatively easy to create a graded series of items spanning a wide range of complexity levels from the simplest kind of counting to the use of calculus (see Table 2). The items are dichotomies asking respondents whether they perform each task as a regular part of their jobs. They avoid using subjective rating scales and are behaviorally specific, while remaining meaningful across occupations. They also avoid the need for respondents to

Table 2
Construct and Criterion Validity of STAMP Math, Reading, and Writing Task Measures

MATH	Item-Rest Correlation	Categorical PCA Loading	Mokken Loevinger H_i	CFA Standardized Loadings	
Counting	0.42	0.49	1.00	1.02	
Add/subtract	0.50	0.57	0.91	0.97	
Multiply/divide	0.53	0.61	0.88	0.94	
Fractions	0.52	0.62	0.88	0.90	
Algebra I	0.63	0.78	0.90		1.01
Algebra II	0.53	0.69	0.79		0.93
Geometry/trig	0.52	0.68	0.71		0.87
Statistics	0.54	0.69	0.72		0.89
Calculus	0.42	0.57	0.83		0.91
	Cronbach's α	Pct. variance	Loevinger H	RMSEA	
	0.81	0.63	0.83	0.06	0.06
<i>Correlations:</i>					
Job education	0.42	0.41	0.41		
Own education	0.32	0.31	0.31		

READING	Item-Rest Correlation	Categorical PCA Loading	Mokken Loevinger H_i	CFA Standardized Loadings	
Any reading	0.30	0.44	1.00	—	
Read one page	0.59	0.75	1.00	—	
Read 5 pages	0.62	0.77	0.64	0.82	
Trade/news articles	0.60	0.74	0.63	0.85	
Prof'l articles	0.62	0.75	0.66	0.89	
Books	0.60	0.75	0.61	0.80	
	Cronbach's α	Pct. variance	Loevinger H	RMSEA	
	0.80	0.75	0.68	0.04	
<i>Correlations</i>					
Job education	0.64	0.63	0.59		
Own education	0.51	0.50	0.46		

WRITING	Item-Rest Correlation	Categorical PCA Loading	Mokken Loevinger H_i	
Any writing	0.30	0.49	1.00	
One page	0.51	0.74	1.00	
Five pages	0.51	0.75	0.83	
Trade/news articles	0.38	0.64	0.70	
Books/prof'l arts.	0.32	0.56	0.79	
	Cronbach's α	Pct. variance	Loevinger H	
	0.64	0.65	0.86	
<i>Correlations</i>				
Job education	0.61	0.60	0.60	
Own education	0.51	0.50	0.50	

rank themselves on a single scale, which might induce self-enhancing biases. The items can be related easily to specific aspects of math curricula, which is relevant to the education researchers and policymakers in the skills mismatch debate. The frequency distribution for these variables provides a relatively objective measure of the levels of math required in the current workplace.

Some test makers and job analyses of narrow occupations use even more detailed and elaborate typologies of math tasks (e.g., single vs. multi-step problems, use of scientific notation), but these schemas are too detailed or complex for self-rating tasks in a short phone survey (Kirsch et al. 1993; Manly et al. 1994; ACT 2002).

The top panel of Table 2 presents various measures of construct validity. The first column gives Cronbach's α for an additive scale of the math items, which is reasonably strong (0.81), and the correlations of each item with a scale excluding that item. Most have item-rest correlations greater than 0.50, indicating strong fit with the other items in the scale.

The second column presents results of a nonlinear principal components analysis, which is more appropriate than standard PCA for dichotomous data (Meulman, Van der Kooij, and Heiser 2004).⁶ The first principal component accounts for 63 percent of the total variance, well above the 30–40 percent cutoff commonly recommended for a strong dominant factor. In results not shown, there is some indication of the utility of identifying a second construct corresponding to basic math use only.

The third column gives results of fitting the items to a Mokken scale, which is a probabilistic Guttman scale. Mokken scales take into account the hierarchical nature of these items, while classical methods assume parallel measures (i.e., equal item means and variances) and are not strictly appropriate for hierarchical sets of items. People who respond positively to higher-level items (e.g., use calculus) are expected to respond positively to all lower-level items (e.g., perform multiplication/division), while the converse is not true by design, which attenuates the inter-item correlations that are the basis of classical methods (Sijtsma and Molenaar 2002, p.55; van Schuur 2003, pp.140f.; Bond and Fox 2001, pp. xiif.).

⁶These calculations were performed using the CatPCA command in SPSS.

Mokken scales use Loevinger's coefficient of homogeneity at the item (H_i) and scale (H) levels to test whether the number of departures from a strictly hierarchical response pattern is low enough that persons and items can be consistently ordered. H_i are measures of item discrimination, somewhat analogous to item-rest correlations in classical item analysis and H is a measure of overall scalability.⁷ These are alternatives to classical measures of a scale's unidimensionality and construct validity.

Values for the full STAMP math scale ($H=0.83$) and each of the items ($H_i >0.70$) are significantly different from zero and show very strong scalability.⁸ Indeed, about 91 percent of all respondents answered the math items in a strictly cumulative fashion, consistent with the Guttman model. These results are highly consistent with the idea that the items represent a single hierarchy of difficulty or skill level.⁹

Results of confirmatory factor analysis (CFA) in the fourth and fifth columns, conducted with Mplus, reinforce the PCA results. A two-factor solution fit significantly better than a single factor solution. Figures in the table are fully standardized loadings and the root mean standard error of approximation (RMSEA) is used as an overall fit index. The fit is acceptable but the hierarchical quality of the items creates an interdependency among the indicators that makes CFA problematic, so these and similar results should be taken as suggestive.

⁷ H_i and H are one minus the ratio of observed to maximum possible Guttman errors, where the latter is the joint distribution of item frequencies based on the item marginals and assuming independence between items. Mokken recommended that all H_i exceed 0.30 and classified H values as indicating weak scalability ($0.30 \leq H \leq 0.40$), medium scalability ($0.40 \leq H \leq 0.50$), and strongly scalability ($H \geq 0.50$). He also developed a test of the null hypotheses that $H_i=0$ and $H=0$ (Sijtsma and Molenaar 2002, p.51ff.; van Schuur 2003, p.149).

⁸These calculations were performed using the `-msp-` command in Stata.

⁹It should be noted that unlike cognitive ability testing, where such models are commonly applied, the assumptions of a cumulative response pattern may be violated in the case of workplace skill demands. The functional division of labor may mean that some people using calculus never have occasion to use geometry for problems in their line of work, for example. In this case, the items represent qualitative variables rather than different levels within a single construct.

Alternatively, the hierarchical division of labor may mean that a more skilled employee or someone on the middle or upper rungs of a career ladder no longer performs more routine tasks that are delegated to lower-level employees, even though the higher-level employee is capable of performing the task and may have done so previously in the normal course of their career (Ludlow 1999, p.973). In this case, models based on observed data might suggest it is not possible to consistently order workers by job skill requirements when theory suggests it is possible.

Two methods of contrasting groups were used to gauge the criterion validity of the items. When the sample is divided into five broad occupation groups, the proportion of positive responses to more difficult math items generally follows a pattern consistent with expectation (results available on request).¹⁰

When contrasting groups are defined by job education requirements (*job education*) and personal education attainment (*own education*), the various math scales correlate significantly with both criteria, and more strongly with job education (≥ 0.40) than with own education (≥ 0.30), further supporting the criterion validity of these job skill measures (Table 2).

2. *Reading and Writing*

Unfortunately, reading matter is not as simple as math to measure on a single continuum and classify into levels that correspond to familiar curricular concepts or years of schooling. Educational psychology has long wrestled with the problem of how to rate the complexity or readability of text in a standardized manner, whether in terms of grade-level equivalents or other units. Readability is itself an ill-defined construct, which makes measuring it through self-reports all the more challenging.

A leading review of readability formulas concluded that the best predictors of reading difficulty level are relatively simple measures of word difficulty (number of syllables, word familiarity), and sentence difficulty (words per sentence) (Klare 1974–1975; Klare 2000).¹¹ In practice, most formulas for measuring a text's complexity level use some combination of its average word length and average sentence length.

Perhaps counterintuitively, more sophisticated measures of word complexity (e.g., number of affixes, Latin-based syllables) and grammatical complexity (e.g., prepositional phrases, subordinate

¹⁰The occupational groups are upper white-collar (managers, professionals, technical workers), lower white-collar (clerical, sales), upper blue-collar (craft, repair workers), lower blue-collar (operators, laborers), and service workers (health, food, personal, and protective service workers).

¹¹Word familiarity is based on whether the word appears on a long list of commonly-used words or the word's frequency of use in a very large, random sample of texts, which are tabulated in lexical compendia or databases (Stenner and Stone 2004).

clauses) have been found to add little to the predictive power of readability formulas. Careful advocates of simple readability formulas do not claim that word and sentence length are the most important causes of text complexity, only that they seem to be very effective proxies or indicators (Klare 1974–1975).

Nevertheless, readability formulas do not capture very rich information about a text's complexity level, such as conceptual complexity, level of reasoning, density of details, amount of distracting information, and level of background knowledge needed to understand the text (Mosenthal 1998; ACT 2002, pp.67ff.). Unfortunately, these dimensions are intrinsically difficult to define and measure across diverse kinds of reading matter. Despite dissatisfactions with readability formulas for their simplicity and focus on surface qualities, no alternative has achieved similar levels of acceptance.

An additional problem is that different readability formulas may correlate with one another, but they do not necessarily agree on the absolute grade-level they assign to the same text, which is necessary for understanding whether persons are well matched to job reading requirements.

Writing at a given level presumably stands in a cumulative relationship to reading at that level. Generating text requires all of the skills needed to interpret it and additional skills, such as the ability to use proper spelling and grammar, knowledge of word use, use of appropriate tone and style, and the ability to logically organize and develop ideas (ACT 2002, pp.19ff.). Ideally, measures of job-related writing tasks would account for the degree to which they required these and other skills.

As the preceding indicates, rating the complexity level of reading and writing tasks is challenging even when researchers have access to the texts in question. There are no relatively clear and widely understood category schemes available to either researchers or survey respondents that capture the complexity involved, as are available for the domain of mathematics.

STAMP addressed this problem with a hierarchy of items based on text length at the low end and text complexity in the middle and upper ranges to integrate the insights of both sides of the debate over readability formulas in education research (middle panel, Table 2). It was hoped that the text descriptors for the middle and upper levels would be well-ordered, relatively unambiguous to respondents, and provide good coverage of the range of text typically encountered in the workplace in terms of both

qualitative variety and difficulty levels. Other reading matter, such as manuals and bills or invoices, were also part of this series but proved qualitatively distinct and did not scale with the other items. Two items designed to measure the highest levels of reading complexity also showed problems.

While most questions asked about reading tasks performed regularly, the item on book reading asked whether respondents ever had to read work-related books as part of their job. The wording is reasonable if book reading is very infrequent but necessary for someone's job. However, it appears to have raised the frequency of positive responses to levels above those for some of the easier items (e.g., reading articles in trade magazines or newspapers), which reverses their expected order of difficulty.

Respondents also may have misinterpreted the meaning of the item on scholarly, scientific, and professional journals, as distinct from the item on trade magazines, newsletters, and newspapers. Despite efforts to distinguish the two for respondents, an unexpectedly large proportion reported reading journals, which may indicate some upward bias. Reassuringly, job and personal education correlate more strongly with reading professional journals (0.54 and 0.44) than trade magazine and news articles (0.49 and 0.38), but the differences are not great.

Unfortunately, given that reading researchers have not developed a standard method for rating texts themselves, it is not surprising that self-reports from laypersons have problems.

Nevertheless, there is evidence for the reading scale's construct validity. The middle panel of Table 2 shows high values for Cronbach's α (0.80), the variance explained in a nonlinear PCA (0.75), and Loevinger's H (0.68). Compared to the math items, fewer respondents answered the reading items in a strictly cumulative fashion without Guttman errors, whether the scale is constructed assuming books are less difficult than both kinds of articles (71 percent) or more difficult (65 percent). Dependencies among the items required estimating a trimmed CFA model, which also showed good fit (RMSEA=0.04). Likewise, most of the measures of item discrimination are quite high.

The reading items also showed high criterion validity. Proportions of positive responses discriminate among the five broad occupational groups as expected (results available from author). The

various scales of reading complexity correlate even more highly with job education (~ 0.60) and personal education (~ 0.50) than the math scales.¹²

The items on job-related writing parallel the reading items except that the items on professional articles and books were collapsed into a single question because few positive responses were anticipated for either item, which was the case.

While Cronbach's α , the nonlinear PCA, and their associated item statistics suggest this scale does not perform as well as the reading scale, the results from the Mokken scale analysis suggest the writing scale performs better. Reflecting this fact, a very high proportion of cases answered this series of items in strictly cumulative fashion (95 percent). The writing items are better than the reading items in discriminating between complexity levels and the proportion of positive responses drops much more sharply after the second level. Perhaps as a result, various CFA models did not fit well and are not reported.

Compared to the reading items, the writing items are more strongly associated with occupational group (not shown) and have similar correlations with job and own education, indicating strong criterion validity (bottom panel, Table 2).

3. *Document Use*

Education researchers have long known that workplace literacy for many jobs involves heavy use of materials such as forms, diagrams, and graphs, which differ from standard, continuous text (Sticht 1975; Mosenthal and Kirsch 1998; ACT 2002, pp.39ff.). Some national tests treat document literacy as a separate dimension of literacy, distinct from prose and quantitative literacy (Kirsch et al 1993). Although

¹²The Mokken scale was constructed assuming that books represented a higher level of difficulty than news or journal articles. Cases with Guttman errors were corrected by arbitrarily assigning values equal to the highest level of self-reported reading task used on the job, with the further restriction that in order for cases with Guttman errors to be coded as reading books the respondents also had to report reading at least one kind of article for their job. The criterion correlations for this scale are somewhat lower than those for the other reading scales. However, if the sample is restricted to cases without Guttman errors the correlations with job education (0.72) and personal education (0.59) would be about 0.10 higher than those using the more conventional scales.

it is not clear that this is the case, the category is potentially important for detecting finer distinctions among people at the very low end of the reading scale.

Mosenthal and Kirsch (1998) coded document complexity based on organizational complexity (e.g., simple lists vs. nested lists) and density (number of organizing categories and items). Fernandez (2001) used this scheme in a case study of the changing skills and wages in a manufacturing plant. However, this method requires that the actual documents are available for analysis by trained coders. STAMP uses several alternative measures for this dimension of job skill demands.

a. Forms. The STAMP question defined this category for respondents by citing examples such as order forms, online forms, bills, invoices, and contracts. Three questions asked respondents the number of different forms they used, the modal length of the forms, and the forms' complexity on a rating scale that varied between 0 and 10. The ends of the rating scale were anchored by examples and labeled extremely simple and extremely complicated, respectively. People who did not use forms on their job were also assigned zero values on all three variables.

Although the inter-correlations among the three items were very high for the entire sample (0.72–0.75), they dropped considerably when calculated for the sub-sample that uses forms, about two-thirds of the total. For this group, form complexity still correlates strongly with form length (0.41) and moderately with the number of forms (0.31), but the number and length of forms were not strongly related to one another (0.22). Clearly, a high proportion of null values can inflate measures of inter-item consistency because the correlations will reflect the fact that none of the items really apply to these cases, rather than a more substantive reason for the associations.

Similarly, Cronbach's α for the restricted sample did not reach conventionally acceptable levels (0.58), but was very high when people not using forms were included in the sample (0.89). The first principal component in a nonlinear PCA explained 67 percent of variance when restricted to form users. However, the scale had only marginally higher criterion correlations with job and personal education than form length and complexity taken individually. Consequently, it remains to be seen whether combining

the three items in a single scale has any predictive advantage over the individual items in substantive models.

b. Visual matter. The STAMP question defined this category for respondents by citing examples that included maps, diagrams, floor plans, graphs, and blueprints. The items asked if respondents used visual matter on their job and if they write or create such materials themselves.

Use of visual materials is weakly correlated with both job and personal education (0.16 and 0.13), but their production is somewhat more strongly related to each (0.24 and 0.19). In retrospect, an item asking length of training required to use or create such documents might have produced stronger measures and criterion validity coefficients.

4. *Problem Solving*

Math, reading, writing, and document use cover a significant portion of the cognitive skills domain and have general applicability to a wide variety of jobs. They correspond closely to school subjects, are the focus of national testing programs for both students and adults, and are related directly to recent policy debates regarding the academic preparedness of students and workers relative to job requirements. However, they do not cover more general thinking and reasoning skills, discussed here, and the multitude of cognitive or technical skills that comprise job-specific human capital, discussed in the next section.

Discussions of the skills crisis often mention the need for problem solving skills, but the term is usually undefined. Problem solving appears to refer to the application of general reasoning ability and common sense to novel or non-routine situations. However, sometimes it appears to include non-cognitive dimensions as well, such as the willingness to go beyond one's immediate duties or supervisor's instructions when encountering unexpected or disruptive situations (conscientiousness), and the ability to cope independently with unpredictable or difficult situations without a supervisor's assistance (Wise, Chia, and Rudner 1990, pp.20,24).

STAMP defined problem solving for respondents as dealing with new or difficult situations that require an employee to think for a while about what to do next; the survey does not elicit information on proactive work orientations. This is consistent with the definition offered by the International Adult Literacy and Lifeskills Survey (ALL) (2003):

Problem solving involves goal-directed thinking and action in situations for which no routine solution procedure is available. The problem solver has a more or less well defined goal, but does not immediately know how to reach it. The incongruence of goals and admissible operators constitutes a problem. The understanding of the problem situation and its step-by-step transformation, based on planning and reasoning, constitute the process of problem solving (Organisation for Economic Co-operation and Development 2005, p.16).

The British Skills Survey also has several questions on problem solving, but they define it in terms of trouble-shooting. This is somewhat narrower than the STAMP and ALL conception, which includes any situation that requires going beyond the information at hand to figure out what action to take next, i.e., problems as puzzles.

STAMP asked respondents how often they faced *easy problems*, defined as those that could be solved right away or after getting a little help from others, and *hard problems*, defined as those that require more time and a lot of work to come up with a solution. The response scales for these questions used vague quantifiers (never, rarely, sometimes, often), but respondents who said they had to solve hard problems on their jobs were then asked the number of hard problems they face in an average week. As a group, they measure both complexity level and frequency, though the level dimension is coarser than for the other variables discussed previously. (For similar items, see van de Ven and Ferry 1980, p.463; Wall, Jackson, Mullarkey 1995, p.439).

The two items dealing with hard problems have high reliability ($\alpha=0.85$) and the scale accounts for a large proportion of the variance in a nonlinear PCA (0.85).¹³ A scale with the two items with vague quantifiers has a much lower reliability ($\alpha=0.70$) and accounts for a smaller proportion of the total

¹³The number of hard problems was logged after adding a very small number to those with zero values and both variables were standardized before the scale was constructed.

variance (0.79), but its correlations with job education (0.45) and personal education (0.36) are about 0.05 higher compared to the first scale. The most behaviorally specific item, the (ln) number of hard problems in an average week, correlates even more weakly with job education (0.30) and personal education (0.23).

In this case, the two less behaviorally specific variables have a stronger relationship to the criterion variables than the more specific item used by itself or as part of a scale. Nevertheless, having a self-report of the number of hard problems is useful because it provides a check on whether respondents within and across samples assign comparable meanings to the vague quantifiers in the first item on the number of hard problems they encounter.

5. *Education, Experience, and Training*

Finally, three STAMP measures of cognitive skill demands are even more global.

1. Level of education needed by the average person to perform the respondent's job (*job education*);
2. Years of prior experience in related jobs needed by someone with that level of education to be qualified for the job (*related job experience*); and
3. Length of time needed by someone with those characteristics to learn how to do the job (*training time*).

All three are measured on objective scales, correspond clearly to concepts in human capital and other theories, and are intuitively meaningful for researchers, policy analysts, practitioners, and laypersons. By design, job education is measured on the same scale as personal education so it is possible to compare them directly and calculate rates of mismatch. By contrast, researchers using the corresponding DOT variable, General Educational Development (GED), had to use their judgment or some estimation procedure to convert GED codes to years of education to make such comparisons (Berg 1971; Halaby 1994).

The experience and training items capture the kind of non-academic and job-specific skills that are otherwise unmeasured because they are too idiosyncratic for a general survey and incommensurate with one another even if it were possible to include lengthy checklists of occupational skills. This is

probably the only way to solve the problem of measuring qualitatively distinct job-specific human capital on a common, absolute scale (Cully et al. 1999, p.63).

Job education and training time also appears in the Panel Study of Income Dynamics. Household heads answered these questions in both 1976 and 1978, permitting calculation of test-retest reliability correlations. The cross-year correlation for job stayers is approximately 0.83 (n=1,356) for job education and 0.60 (n=1,446) for training times.¹⁴ By contrast, the equivalent correlations for those who changed employers and three-digit occupation and industry were 0.51 (n=228) for job education and 0.22 (n=257) for training times (Handel 2000, pp.187f.). Thus, there is a reasonable level of consistency for job stayers, particularly for job education, and considerably less consistency for job changers, as expected.

The pattern of correlations suggests respondents may be using their own educational level partly as a guide in reporting the education required for their job, which would bias responses upward. However, the fact that almost all other skill measures are more strongly associated with job education than personal education suggests that the measure contains unique information and has strong criterion validity.

Once the second wave of the STAMP survey is completed, similar test-retest reliability correlations will be available for all items repeated across the two waves.

6. *Interpersonal Skills*

Job-related interpersonal skills remain even less well-theorized and well-measured than cognitive skill requirements, despite the growing research literature. Even at the most basic level, this domain is weakly conceptualized. This makes it difficult to assess content and construct validity because judging how thoroughly or consistently a domain has been measured requires knowing what belongs in it.

The International Adult Literacy and Lifeskills Survey (2003) tried to measure teamwork, which was defined as the skills needed to function in any work group in which tasks require interpersonal coordination and communication and group decision making. The project found little research on the

¹⁴Job stayers were defined as respondents with job tenure equal to or greater than two years in 1978 and whose reported three-digit occupation and industry were identical across the 1976 and 1978 interviews.

topic, had difficulty defining different levels of team skills, and abandoned this part of the planned assessment after failing to design acceptable measures (Statistics Canada 2005, pp.228ff.).

ACT's WorkKeys test has a section on teamwork skills that covers cooperativeness, positive attitudes, interpersonal trust, attention to others' input, time management, goal direction, planning, problem solving, creativity, conscientiousness, leadership, and delegation, among others (ACT 2002, pp.79ff.). Test-takers respond to questions based on video vignettes. However, employers show much less interest in this subtest than the sections on math and reading (ACT 2002, p.59).

From these and other sources, it appears that interpersonal skills include communication skills, courtesy and friendliness, service orientation, caring, empathy, counseling, selling skills, persuasion and negotiation, and, less commonly, assertiveness, aggressiveness, and even hostility, at least in relations with organizational outsiders that are adversarial (e.g., police and corrections officers, security guards, bill collectors, some lawyers and businessmen, etc.) (Hochschild 1983; U.S. Department of Labor 1991). If one were to include informal job demands that might arise in dealing with co-workers, this list would be expanded to include leadership, cooperation, teamwork skills, and mentoring skills.

Defining the domain of interpersonal job requirements in this way raises several problems. The elements are qualitatively diverse, rather than clearly different levels of a single trait. Many might be considered ancillary job characteristics, which, while often useful in the workplace, are exercised at the discretion of the employee, rather than job or employer requirements. It is often not easy to separate interpersonal skills from more purely attitudinal and motivational aspects of work orientations (Moss and Tilly 2001).

On a practical level, very general questions about job-related interpersonal demands are likely to produce high rates of agreement and low variance if they did not distinguish situations related to co-workers from those relating to contact with customers, clients, and other organizational outsiders. Survey pretests showed a tendency for many people to respond reflexively that working always requires a positive attitude, willingness to cooperate with others, etc. Previous research also shows the pressures on most managers to engage in intensive impression management to maintain smooth working relations with

others within their organization (Kanter 1973; Jackall 1988; Morrill 1995; cf. Reisman, Glazer, and Denney 1950 on other-directedness).

To reduce yea-saying biases, STAMP asked a set of relatively specific items that could apply to relations with either organizational insiders or outsiders and more general questions about extended interactions with outsiders only.

STAMP asks respondents if their jobs require them to give people information, counsel people, deal with tense or hostile people, teach or train people, interview people, and give formal presentations lasting at least 15 minutes. They are also asked if they have contact with customers or the public, the frequency of contacts lasting more than 15 minutes, and self-rated importance of such contact for their jobs. While these items attempt to be relatively concrete, substantial room for individual interpretation and yea-saying biases undoubtedly remains.

STAMP does not ask about caring labor because pretests suggested it has salience only for people who work in jobs that could be identified easily from intuition, existing research, and occupational title alone. The same consideration applied to selling skills. The more generic negotiating, persuading, and influencing skills had the reverse problem. Rather than applying too narrowly, they seemed too open to interpretation and overly positive responses to warrant using scarce time (U.S. Department of Labor 1991, pp.3–6ff.). Obviously, there are arguments for making opposite choices, as well.

A scale using the measures of interpersonal skills has reasonable values for Cronbach's α (0.72), variance explained (0.68), and CFA model fit (RMSEA=0.04), and for the corresponding measures of item discrimination (Table 3).

A number of items have relatively high levels of endorsement. However, the more specific items on working with the public yield a stronger gradient across occupations in the expected manner (not shown).¹⁵ Correlations with both educational measures are within the range found for the cognitive skills

¹⁵On an 11-point importance scale (11=extremely important), the value for lower blue-collar workers is 4.2, upper blue-collar workers is 5.0, service workers is 6.9, lower white-collar workers is 8.3, and upper white-collar workers is 8.8.

Table 3
Criterion and Construct Validity of STAMP Interpersonal and Physical Task Measures

INTERPERSONAL	Item-Rest Correlation	Categorical PCA Loading	CFA Standard Loading
Information	0.30	0.44	0.81
Counseling	0.43	0.54	0.66
Tense situations	0.34	0.44	0.40
Teach/train	0.37	0.50	0.66
Interview	0.38	0.53	0.74
Presentations	0.43	0.57	0.75
Public contact ^a	0.54	0.80	0.44
Self-rated level ^b	0.54	0.80	0.46
	Cronbach's α	Pct. variance	RMSEA
	0.72	0.68	0.04
<i>Correlations</i>			
Job education	0.48	0.48	
Own education	0.39	0.40	
Management	0.40	0.40	
PHYSICAL	Item-Rest Correlation	Categorical PCA Loading	CFA Standard Loading
Stand 2 hours	0.56	0.76	0.82
Lift 50 lbs.	0.57	0.77	0.85
Coordination	0.53	0.73	0.74
Physical demands ^c	0.72	0.87	0.84
	Cronbach's α	Pct. variance	RMSEA
	0.79	0.80	0.00
<i>Correlations</i>			
Job education	-0.33	-0.34	
Own education	-0.34	-0.34	

Note: The CFA model for interpersonal skills allows “tense situations” to correlate with “counseling” and frequency of contact with the public (“public contact”) and allows “public contact” to correlate with importance of working well with the public (“Self-rated level”).

^aSix-point scale measuring frequency of contact with people other than co-workers, such as customers, clients, students, or the public (0=none, 5=spend at least 15 minutes speaking to non-coworker more than once a day).

^bSelf-rated importance of working well with customers, clients, students, or the public on respondent’s job (0=not important, 11=extremely important).

^cSelf-rated physical demands of job (0=not all physically demanding, 10=extremely physically demanding)

scales, though it is not clear whether either education variable is suitable for establishing criterion validity in the case of interpersonal skills (Table 3). Table 3 also includes the correlation with manager status as another way to measure the criterion validity of the interpersonal skills scale, but this also must be considered exploratory. There is little existing research on the criterion validity of interpersonal job task measures.

7. *Physical Demands*

Physical job requirements are bodily actions that usually involve materials, tools, and equipment and represent the final category of the DOT's classification of job tasks into work with Data, People, and Things.

Examples of less skilled physical tasks include simple physical exertion (e.g., carrying heavy loads), simple movements (e.g., sorting mail), use of simple tools or equipment, or monitoring and tending machines or other physical processes. Higher skilled tasks often require more training, experience, and background knowledge regarding the properties of physical materials, mechanical processes, and natural laws (U.S. Department of Labor 1991, pp.3–11 ff., 12–1 ff). Complex physical skills are undoubtedly more difficult to capture using a limited number of items, which means that many craft skills, which are often occupation-specific, remain unmeasured.

The STAMP items in the bottom panel of Table 3 are self-explanatory, except perhaps the item on hand-eye coordination and arm steadiness, which was the lone attempt to capture more skilled physical demands. The items on standing and lifting follow the preferred approach of measuring job characteristics using objective yardsticks. The last STAMP item is a more global self-rating of jobs' physical demands using a subjective rating scale for capturing otherwise unmeasured characteristics.

This set of questions is relatively short because previous research generally shows that the physical job tasks are declining relative to cognitive and interpersonal tasks (Bell 1973; Zuboff 1988; Reich 1991; Handel 2000). In addition, physical requirements have generally weak effects on wages and

are not prominent in recent debates over work and employment (Brown 1980; Rotundo and Sackett 2004).

Scales composed of the four items have strong construct validity ($\alpha=0.79$, variance explained=0.80, RMSEA=0.00) (Table 3). Blue-collar and service workers are much more likely to report their jobs involve physical work. Skilled blue-collar workers are the most likely to say their work requires good eye-hand coordination or a steady hand. The scales have good criterion validity, correlating negatively with job and personal education (~ 0.34).

B. Technology

There have been many efforts to develop standardized measures of technology, both nominal classifications and ordinal scales. Many are useful, but none has achieved widespread acceptance.¹⁶ Despite its centrality to research on skill, work, and organizations, technology remains an under-theorized domain and its salient dimensions relatively undefined.

The concept “computer literacy,” for example, has surprisingly little precise meaning despite its common use by both professionals and the general public (Goodson and Mangan 1996). Psychologists have spent more effort on scales measuring attitudes, such as computer anxiety and self-efficacy, rather than task complexity (Loyd and Gressard 1984; Heinssen, Glass, and Knight 1988; Bunz 2004, pp.483f.). ACT’s (2002) WorkKeys assessment has no test of computer literacy. The International Adult Literacy and Lifeskills Survey (ALL) (2003) tried but failed to develop reliable measures of computer skills and task complexity. The final instrument had items on attitudes, familiarity, and use of mostly internet-related applications, which arguably are not high on the complexity scale (Statistics Canada 2005, p.23; Organisation for Economic Co-operation and Development 2005, 186ff.).

¹⁶For notable examples, see Woodward 1965; Blau et al. 1976; Hull and Collins 1987; Schmenner 1987; Kalleberg and Lincoln 1988; Krueger 1993; Doms, Dunne, and Troske. 1997; Liker, Haddad, and Karlin 1999; Melcher, Khouja, and Booth 2002.

STAMP has fifty questions on computers, automation, and non-computer technology that attempt to capture the incidence of the most important and prevalent workplace technologies and the skill levels they require.

1. *Information Technology*

STAMP addressed the issue of measuring task complexity using four approaches.

The first was to ask eighteen questions about the use of specific computer applications, expanding the checklist approach used by the Current Population Survey. These items were used to construct a count variable for the number of software applications used on the job. STAMP also asks the frequency of computer use on the job.

In an effort to move away from a strictly checklist approach and measure complexity more directly, STAMP also included five items designed to capture higher-level computer tasks (e.g., scientific/engineering calculations), as well as one highly routine activity (data entry) that has been the focus of deskilling approaches (Braverman 1974; Hartmann, Kraut, and Tilly 1986–1987). Although a natural extension of the checklist approach, few surveys include such items.

A third approach that departs more significantly from previous efforts measures the length of computer learning times, which is a natural unit for measuring skill in human capital theory. Most people would have difficulty remembering how they learned general computer skills, because the process is gradual and continuous rather than a discrete period. Therefore, STAMP asked if respondents used computer programs that were specific to their job and, if so, how long it took them to learn the most complex such program. If they had to learn any new computer program on their job in the previous three years, they were asked also how long it took to learn the most complex such program. These appear to be the first attempts to produce numerical estimates of the complexity of computer applications used on the job. The second item is also the first measure of the rate of technological change experienced by workers.

The fourth approach is a measure of overall computer task complexity using a subjective rating scale varying from “very basic” (0) to “very complex” (10).

Finally, to measure possible computer skill deficits, STAMP asked respondents if they had all the computer skills they needed for their current job and if lack of computer skills had affected their chances of employment, promotion, or pay raise.

A scale combining a count of software applications used on the job and self-rated complexity of computer tasks has a reasonable Cronbach's α (0.71) and the first principal component accounts for a large share of total variance (0.77) (Table 4). The scale is more highly correlated with job educational requirements (0.43) than own education (0.31), providing evidence of criterion validity. These calculations are conservative because they use only the sub-sample of computer users. Using the whole sample increases the scale's α (0.89) and its correlations with job education (0.56) and own education (0.45) (not shown).

Using the method of contrasting groups, one finds clear differences consistent with intuition in both the number of software applications used and task complexity across broad occupational groups. Upper white-collar workers used 6 applications on average, while service workers used less than 1.5. Large proportions of white-collar workers use special software (~60 percent), while this was much less common for blue-collar and service workers (~25 percent). Upper white-collar workers were more likely than other groups to have had to learn new software in the previous three years. A similar pattern held for other higher-level tasks, such as the use of spreadsheet macros and formulas.

The two items regarding computer skill deficits did not scale with the other items or with one another ($\alpha=0.30$). Their correlations with both measures of education are less than 0.10 (not shown).

2. *Non-Computer Technology*

While office computers have received the most attention, the technology domain also includes non-computer technology such as heavy machinery and industrial equipment, which are even less commonly covered in other surveys.

Various tests of mechanical and other technology skills help define this domain. The WorkKeys' Applied Technology sub-test covers basic physical forces, mechanical systems, electricity, plumbing,

Table 4
Criterion and Construct Validity of STAMP Computer Technology and
Employee Involvement (EI) Measures

COMPUTERS	Item-Scale^b		
	Correlation	PCA Loading	
Applications (#)	0.88	0.71	
Skill Level ^a	0.88	0.71	
	Cronbach's α	Pct. variance	
	0.71	0.77	
<i>Correlations</i>			
Job education	0.43 (0.56)	0.44 (0.56)	
Own education	0.31 (0.45)	0.32 (0.46)	
EI^c	Item-Scale	Categorical	CFA Standard
	Correlation	PCA Loading	Loading
Job assignment	0.38	0.64	0.59
Task scheduling	0.44	0.73	0.67
Worker scheduling	0.36	0.58	0.54
Changing methods	0.42	0.69	0.60
New equipment	0.40	0.60	0.51
Selecting leader	0.19	0.20	0.26
Quality	0.42	0.52	0.65
Cost, productivity	0.43	0.53	0.65
Cross-communicate	0.31	0.42	0.48
Performance review	0.19	0.24	0.29
	Cronbach's α	Pct. variance	RMSEA
	0.69	0.67	0.09
<i>Correlations</i>			
Decision making	0.16		
Autonomy scale	0.14		
Job education	0.10		
Own education	0.13		

Note: Statistics in top panel calculated for computer users only, except those in parentheses, which are based on all respondents.

^aSelf-rated complexity of computer skills used on job (0=very basic, 10=very complex)

^bCorrelations of items with the total scale are presented instead of item-rest correlations because the latter is identical for both rows and the correlation between the two items themselves, which is 0.55.

^cStatistics based on sub-sample belonging to teams. Employees in self-reported management positions were ineligible for these items.

hydraulics, pneumatics, and heating and refrigeration systems (ACT 2002, p.9). This is a good map of the content of this domain, but it is stronger on traditional craft skills than newer technical skills, omits deskilled technologies, and is more detailed than possible in a short, general-purpose survey.

STAMP respondents who used heavy machinery on their jobs other than vehicles were asked sixteen items covering social science concerns regarding traditional craft skills (machine set-up, maintenance, repair), newer high-technology skills (programmable automation technology), and deskilled tasks (machine tending, assembly line work).

Respondents were also asked whether any new equipment was introduced in the previous three years and the time required to learn the most complex new equipment, providing numerical estimates of both the skill requirements of recent technology and rates of technological change.

In addition, all respondents were asked to rate the level of mechanical knowledge needed for their jobs (0–10 scale) and whether they needed “a good knowledge of electronics, such as understanding transistors or circuits,” for their jobs (1=yes).

Finally, respondents reporting they lost their job in the previous three years were asked if it was because a machine or computer had replaced them to generate estimates of technological displacement.¹⁷

Intuition suggests that occupational group is the most meaningful criterion variable for validating these items. Consistent with expectation, the groups most likely to use heavy machinery and industrial equipment are upper blue-collar (65 percent) and lower blue-collar (46 percent) workers, while the other three occupational groups have very low incidence rates (~10 percent). The ratings for required level of mechanical knowledge exhibit a similar pattern.

¹⁷A small sub-sample of non-employed respondents who had worked at any time in the previous three years was asked the same question to ensure full coverage of the population at risk.

C. Management Practices

Management practices are the third principal leg of recent debates over work and employment. This includes employee involvement practices, other aspects of autonomy and control in the workplace, and various downgrading practices emphasized by the lean and mean approach.

1. *Employee Involvement (EI)*

The major elements of employee involvement programs are relatively well defined: job rotation/cross-training, formal quality programs, self-directed teams, and supportive training and compensation policies (e.g., pay for skill, gain-sharing, performance-based pay). STAMP covers all of these dimensions but the focus here is on teams in the interest of conserving space.

Some previous surveys used very general questions regarding teams and quality, which research suggests may evoke erroneous positive responses from employees in traditional workplaces who understand the terms in more commonsense fashion (Cully et al. 1999, p.42ff.).

STAMP's EI questions are as specific and concrete as possible to minimize ambiguity and false positives. Thus, questions on team membership probe for the frequency of meetings and include ten questions on specific areas of team authority and responsibilities that correspond to the use the team concept in the research literature, adapted partly from a well-known multi-industry survey of employee involvement (Appelbaum, Bailey, Berg, and Kalleberg 2000).¹⁸

The ten team authority items have adequate construct validity using only the sub-sample of team members for conservative tests ($\alpha=0.69$, variance explained=0.67, RMSEA=0.09) (Table 4). These items do not form a consistent hierarchy according to a Mokken scaling procedure (not shown).

Although the literature generally implies that EI is more common among less skilled, particularly blue-collar workers, there is no strong association between the EI measures and occupational group (not

¹⁸I thank Eileen Appelbaum for providing me with the employee questionnaires used in Appelbaum et al. (2000).

shown). This may reflect relatively uniform diffusion or problems with the questions; the state of current understanding regarding these practices is still too limited to draw strong conclusions.

Other criterion validity coefficients for the team scale are rather low. Among those who belong to self-directed teams, the team authority scale correlates weakly with perceptions of input into policy making, a four-item autonomy scale, job education, and own education, though it is not clear that the latter two are suitable criteria for team authority (Table 4). The corresponding correlations using the full sample are even lower (not shown).

Despite trying to achieve precision, the STAMP EI items probably contain significant noise, consistent with other research (Gerhart et al. 2000).

2. *Autonomy, Closeness of Supervision, and Authority*

Several interrelated concepts central to the sociology of work are related to EI and skills, but also distinct: autonomy-control, closeness of supervision, and workplace authority.

Autonomy refers to discretion and self-direction, as opposed to external control (Spenner 1990; Green and James 2003). Employees with high autonomy work independently and can initiate new activities and work procedures (Friedman 1999, p.60). Some distinguish between having the ability to set goals and basic rules, and the ability to determine how to meet goals set by others within an existing structure. Psychologists call the former control over a work situation, or strategic autonomy, and the latter control within a work situation, or operational autonomy (de Jonge 1995, pp.25f.). Braverman's criticism of EI practices relative to his view of craft autonomy relies implicitly on a similar distinction (Braverman 1974, pp.35ff.).

Most measures focus on operational autonomy, such as control over hours of work and break times and the sequence and pacing of job tasks, closeness of supervision, restrictiveness of rules, and

level of task routinization. In contrast, strategic autonomy includes authority to make decisions, direct subordinates, and allocate resources.¹⁹

There is no objective standard for many of these concepts and jobs are so qualitatively diverse that it is hard to generate items measuring autonomy that are both concrete and widely applicable. Indefinite questions are vulnerable to people rating their jobs relative to similar jobs, rather than on a single yardstick that spans the full range of job autonomy-control, and therefore more open to self-enhancing biases. Indeed, previous research finds people tend to give themselves relatively high ratings on autonomy and control, while managers give their employees lower ratings (Cully et al. 1999, pp.40ff.; Handel 2000; Gallie et al. 2004, p.249; White et al. 2004, pp.132f.).

Erik Olin Wright conducted one of the most sophisticated efforts to measure autonomy in a large sample survey, asking respondents to describe a case in which they implemented their own ideas on the job, which was then coded into levels by two independent raters (Wright 1985, pp.314ff.). However, even he concluded, “Autonomy within the labor process proves to be an extremely elusive concept...all operationalizations seem relatively unreliable” (Wright 1997, p.55).²⁰

STAMP contains four items on freedom from prescriptive rules and supervisor’s instructions, task repetitiveness, closeness of supervision, and policy-making authority, covering both strategic and operational autonomy. The uncertain payoffs to further items argued for such a short series.

Measures of construct validity for this set of items are very mixed (Table 5). Cronbach’s α is quite low (0.49), albeit similar to some other autonomy scales (Kalleberg and Lincoln 1988, p.S136), and

¹⁹For examples see Brown 1969, p.350; Cook et al. 1981, pp.183ff.; Kohn and Schooler 1983, pp.22ff; Wright 1985, p.316; Breaugh and Becker 1987; Kalleberg and Lincoln 1988; de Jonge 1995, p.65; Wall et al. 1995, p.439; Bosma et al. 1997; Cully et al. 1999, pp.141f; Fields 2002; Burr et al. 2003, p.273; Gallie, Felstead, and Green 2004, pp.247ff.; McGovern, Smeaton, and Hill 2004.

²⁰The picture is not necessarily quite so bleak. Hackman and Oldham’s Job Diagnostic Survey (JDS) is probably the most widely used single instrument among psychologists. A meta-analysis suggested Cronbach’s α for the autonomy sub-scale is 0.69 and the test-retest correlation is 0.63; other studies found self-reports of autonomy correlate with supervisors’ or analysts’ ratings between 0.20–0.45 (Breaugh 1998; Spector and Fox 2003, p.419; Liu, Spector, and Jex 2005, p.331; see also Gerhart 1988, p.157). The autonomy scale used in the *Working in Britain in 2000* survey had Cronbach’s α of 0.66 (McGovern, Smeaton, and Hill 2004, p.246). Nevertheless, few would consider these figures strong evidence of reliability.

Table 5
Criterion and Construct Validity of STAMP Autonomy and Supervisory Task Measures

AUTONOMY	Item-Rest Correlation	Categorical PCA Loading		
Autonomy	0.33	0.71		
Repetitiveness	-0.26	-0.59		
Close supervision	-0.27	-0.57		
Decision making	0.30	0.67		
	Cronbach's α	Pct. variance		
	0.49	0.64		
<i>Correlations</i>				
Job education	0.43	0.42		
Own education	0.34	0.33		
Management	0.39	0.40		
Problem-solving	0.25	0.22		
Job satisfaction	0.29	0.28		
SUPERVISORY	Item-Rest Correlation	Categorical PCA Loading	Mokken Scale Loevinger H_i	CFA Standard Loading
Work assignments	0.46	0.61	0.72	0.74
Discipline	0.71	0.84	0.71	0.92
Performance review	0.67	0.81	0.69	0.88
Pay and promotion	0.69	0.82	0.71	0.89
Hire and fire	0.67	0.81	0.82	0.93
	Cronbach's α	Pct. variance	Loevinger H	RMSEA
	0.84	0.81	0.73	0.03
<i>Correlations</i>				
Management	0.64	0.64	0.70	
Job education	0.31	0.31	0.33	
Own education	0.24	0.24	0.27	

Note: Problem-solving is an additive scale of two items dealing with the frequency of solving easy and complex problems. Figures in the bottom panel were calculated using the sub-sample reporting supervisory responsibilities over others. The correlations for the Mokken scales in the bottom panel used only cases with response patterns that were strictly cumulative (77% of cases).

the measures of item discrimination are weak. By contrast, the variance explained by the first component of a nonlinear PCA is respectable (0.64) and the item loadings are high.

[The criterion validity of the autonomy items is also mixed. The percentage of upper white-collar workers saying they participate in policy making decisions is far larger (44 percent) than the figures for lower white-collar workers and upper blue-collar workers (~20 percent), which are somewhat higher than the percentages for lower blue-collar and service workers (~15 percent) (not shown). Occupational differences for the other three autonomy items are much smaller. However, the additive scale composed of all four items discriminates among occupations reasonably well.

The scale also correlates relatively well with job education (0.43), own education (0.34), and management status (0.39), but not as well with other criteria, such as problem solving (0.25) and job satisfaction (0.29) (Table 5). The scale's correlation with the EI scale (0.15) is even lower, but this also reflects generally weak correlations for the EI scale.

From this evidence, it seems most reasonable to consider that these variables form an index rather than a scale. Indeed, as described above, the domain spans a wide range of concepts rather than representing a single latent trait in any obvious way.

STAMP also contains five items on supervisory authority over others, covering both task and sanctioning authority (Wright 1985, pp.303ff.). In keeping with explicit scaling principles, these items are more concrete and correspond to the principal criteria for defining supervisory jobs in U.S. labor law (Freidson 1986, pp.136ff.).

A scale composed of these items shows strong construct validity and has a hierarchical, Guttman-scale property (Cronbach's α =0.84, variance explained=0.81, RMSEA=0.03, Loevinger's H=0.73) (bottom panel, Table 5). The scale's strong association with managerial and supervisory positions also shows high criterion validity.

3. *Job Downgrading*

While standard data sets contain many relevant variables for testing the job downgrading thesis that originated with Bluestone and Harrison's work, others are absent or weakly represented. Concepts like downsizing, outsourcing, and wage, benefit givebacks are amenable to direct, objective questions, while others, like effort levels and stress, are more challenging.

a) Employment, Pay, and Promotions. For downsizing and particularly outsourcing, STAMP asked respondents whether, in the past three years, they were laid off personally, their workplace laid off a large number of employees, and a significant part of the work usually done at their workplace had been transferred elsewhere. Employees' knowledge of the latter is likely limited, so the item undoubtedly generates lower bound estimates, but almost no other measures of this important phenomenon exist.

STAMP also asked respondents if their employer reduced their pay in the previous three years and, if so, the size of the pay cut and whether their employer cited affordability as the reason. Respondents were also asked if their employer reduced contributions to their retirement and health plans in the previous three years if they received these benefits.

Internal labor markets (ILMs) distinguishing positions with promotion potential from more "dead-end" jobs has been a central concept in the sociology of labor markets but consensus on its operationalization surprisingly elusive (Althausser 1989). Clean measures that avoid contamination from other concepts are difficult to construct.

Idiosyncratic jobs may lack formal job ladders but still offer good opportunities for internal advancement. Small organizations, such as high-tech startups, may not even have opportunities for internal advancement but provide skills and experience valuable for external advancement. Individuals might have few opportunities for promotion if their career has plateaued or they are already so high in the organizational pyramid that few higher jobs exist. Promotion prospects may be affected more by rates of organizational growth or decline and rates of turnover than employer promotion policies.

STAMP asked respondents how likely they were to be promoted in the following three years, similar to Kalleberg and Lincoln (1988, p.S138), and how satisfied respondents were with their promotion

opportunities. The second wave adds items focused more directly on employer practices by asking how commonly their organization promotes people holding similar jobs and contingent employment status (temp, on-call employee, independent contractor, consultant, seasonal worker).

b) Work Effort and Job Stress. Even more difficult is devising an absolute and objective measure of work effort intensity and related job stress, which is part of the political economy argument regarding “lean and mean” management strategies. The fields of IO psychology, occupational health, human factors, and ergonomics have mapped this domain into a large number of highly diverse facets and indicators. These include physical and mental work load, rapid or hectic work pace, tight deadlines, work backlogs, long or irregular hours, infrequent rest breaks, threat of job loss, ambiguous or conflicting demands, heavy responsibilities, emotional labor, and risk of physical harm on the job.²¹

These diverse concepts illustrate the difficulties defining the constructs and developing objective or standard measures applicable across occupations. Indeed, even in controlled, experimental conditions, physiological and work sample measures do not always correlate highly with self-reports or with one another (Hart and Staveland 1988).

When people evaluate the workload of a task there is no objective standard (e.g., its “actual” workload) against which their evaluations can be compared. In addition there are no physical units of measurement that are appropriate for quantifying workload or many of its component attributes... There is no objective workload continuum, the “zero” point and upper limits are unclear, and the intervals are often arbitrarily assigned (Hart and Staveland 1988, p.143).

A more recent study said simply, “One of the most striking features of the whole literature on job strain/job demands is the almost complete lack of clear conceptual definitions” of relevant constructs despite 25 years of research (Kristensen et al. 2004, p.306; see also Kasl 1998, Annett 2002).

Because the political economy approach argues that work has become more intense over time, STAMP asked respondents if their workload increased in the last three years. To help identify whether

²¹For details, see Department of Health and Human Services. National Institute for Occupational Safety and Health n.d., p.9; Karasek 1979; Cook et al. 1981, pp.204f; Kristensen et al. 2004; Siegrist et al. 2004; European Foundation for the Improvement of Living and Working Conditions 2006.

this reflected a management understaffing strategy, those responding positively were asked if they had to shoulder additional tasks because other workers no longer perform them. Two other items asked if their work pace and job stress levels had changed over the last three years (increased, decreased, stayed the same). As anticipated, these retrospective measures show suspiciously high rates of positive responses. This may reflect actual work intensification, but possibly results of career progression, distorted cognitions, and yea-saying bias.

Green (2006) makes a compelling case for avoiding retrospective measures of work effort because of their potential biases and comparing level scores across multiple survey waves. This relies on the implicit and somewhat problematic assumption that the meaning of the response categories remains stable over time, i.e., people do not adapt to new situations in ways that lead them to adjust their standards for what constitutes light or heavy effort or stress over time. The second wave of STAMP will incorporate level measures, which could be used to construct first-difference measures in the event of a third survey wave.

D. Further Validity Measures

As a final validity test, the scales described in previous sections are compared to one another. When variables that are expected to inter-correlate do so, they demonstrate convergent validity, and when unrelated variables fail to correlate strongly they demonstrate divergent validity (Anastasi 1982).

Table 6 presents the correlations among all the STAMP additive scales discussed above. The various measures of cognitive skills often have correlations with one another and with autonomy and computer use greater than 0.40, though some of the largest correlations in the table are between the interpersonal skills scale and the reading and writing scales. The physical demands scale is unrelated or somewhat negatively related to most of the other variables, as expected. Computer use has among the strongest positive associations with the cognitive skills variables and negative associations with the physical demands scale, as expected since computers are most suited to white-collar job tasks. Interestingly, the team scales are not strongly related to any of the other variables, including the

Table 6
Correlations among STAMP Scales

	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.	14.
1. Math	1.00													
2. Read	0.44	1.00												
3. Write	0.42	0.65	1.00											
4. Form	0.25	0.33	0.34	1.00										
5. Form2	0.39	0.47	0.45	1.00	1.00									
6. Visual	0.42	0.33	0.31	0.15	0.24	1.00								
7. Problem	0.39	0.49	0.44	0.23	0.32	0.30	1.00							
8. Problem #	0.30	0.36	0.32	0.19	0.28	0.21	0.73	1.00						
9. Interaction	0.34	0.58	0.51	0.31	0.42	0.24	0.49	0.37	1.00					
10. Physical	-0.07	-0.20	-0.30	-0.13	-0.20	0.04	-0.10	-0.06	-0.09	1.00				
11. Autonomy	0.28	0.41	0.45	0.21	0.31	0.22	0.24	0.16	0.36	-0.28	1.00			
12. Supervise	0.21	0.31	0.34	0.22	0.33	0.13	0.21	0.15	0.50	-0.22	0.39	1.00		
13. Computer	0.42	0.43	0.48	0.37	0.41	0.30	0.37	0.27	0.27	-0.32	0.36	0.26	1.00	
14. Computer2	0.45	0.55	0.56	0.40	0.50	0.27	0.43	0.32	0.44	-0.44	0.36	0.33	1.00	1.00
15. Teams	0.23	0.15	0.15	0.10	0.18	0.19	0.21	0.15	0.23	0.05	0.15	0.23	0.18	0.21
16. Teams2	0.10	0.16	0.09	0.03	0.06	0.13	0.13	0.12	0.14	0.10	0.01	-0.08	0.06	0.08

Note: Form2, Computer2, and Teams2 were constructed for the entire sample, rather than only those eligible for the items, and ineligible respondents were given zero values. Problem # refers to the (ln) number of hard problems respondents faced in an average week. Correlation values greater than |0.40| are in bold.

interaction scale, whether restricting the sample to team members (Teams) or using the whole sample (Teams2).

In general, constructs correlate with others for which relationships are expected and show weak associations with others for which relationships are not expected, indicating both convergent and divergent validity.

IV. CONCLUSION

In the last twenty years, there has been growing interest in skills, technology, and management practices and their social implications. However, understanding them requires studying them directly rather than relying on inferences based on indirect measures collected for other purposes. The various debates over inequality and job upgrading/downgrading will always be constrained by speculation in the absence of surveys that ask employees about specific job characteristics in concrete detail, as many now recognize.

The survey of Skills, Technology, and Management Practices tries to address the knowledge gap by improving upon past measurement practice, which has not advanced greatly since the Dictionary of Occupational Titles and Quality of Employment Survey were conducted in the 1970s. STAMP attempts to define content domains and constructs in a systematic manner, guided by theory. Adopting an explicit scaling approach, most measures are objective, behaviorally concrete items and scales with natural units, rather than rating scales and vague quantifiers, where possible. Items and scales are problem-relevant, clear, and made to be as interpretable as possible, not only for academic researchers across different disciplines but also for policy professionals, practitioners, and the general public. They help satisfy a key interest among all these constituencies, which is the measurement of job characteristics in terms that can be related to measures and common-sense understandings of person characteristics, in order to assess the congruence or mismatch between people and jobs. The effort generally succeeds in producing survey items with high face, content, criterion, and construct validity.

As noted, limitations remain, some of which are more easily addressed than others. The reading items for the higher levels of complexity do not function as well as hoped, resulting in probable ceiling effects. Several domains, such as interpersonal skills, job autonomy, and work effort, are not well-defined in the literature and therefore the measures likely contain greater error variance.

The usefulness of the measures in answering important substantive questions is the ultimate test of their quality. If they are effective, they will be useful to replicate in the future on a regular basis as social indicators for monitoring trends in job content and related job characteristics (Land 1983).

References

- Advisory Panel for the Dictionary of Occupational Titles (APDOT). 1993. "The New DOT: A Database of Occupational Titles for the Twenty-First Century (Final Report)." Washington, DC: U.S. Department of Labor
- ACT. 2002. *WorkKeys® Test Descriptions*. Iowa City, IA: ACT.
- Althausen, Robert P. 1989. "Internal Labor Markets." *Annual Review of Sociology* 15: 143–161.
- Anastasi, Anne. 1982. *Psychological Testing*. New York: Macmillan.
- Annett, John. 2002. "Subjective Rating Scales: Science or Art?" *Ergonomics* 45: 966–987.
- Appelbaum, Eileen, and Rosemary Batt. 1994. *The New American Workplace*. Ithaca, NY: ILR Press.
- Appelbaum, Eileen, Thomas Bailey, Peter Berg, and Arne L. Kalleberg. 2000. *Manufacturing Advantage: Why High-Performance Work Systems Pay Off*. Ithaca, NY: ILR Press, 2000.
- Attewell, Paul. 1987. "The Deskilling Controversy." *Work and Occupations* 14: 323–46.
- Attewell, Paul. 1989. "The Clerk Deskilled: A Study in False Nostalgia." *Journal of Historical Sociology* 2: 357–388.
- Attewell, Paul. 1990. "What is Skill?" *Work and Occupations* 17: 422–447.
- Autor, David H., Lawrence F. Katz, and Alan B. Krueger. 1998. "Computing Inequality: Have Computers Changed the Labor Market?" *Quarterly Journal of Economics* 113: 1169–1213.
- Autor, David H., Frank Levy, Richard J. Murnane. 2002. "Upstairs, Downstairs: Computers and Skills on Two Floors of a Large Bank." *Industrial and Labor Relations Review* 55: 432–447.
- Barton, Paul E. 2000. "What Jobs Require: Literacy, Education, and Training, 1940–2006." Princeton, NJ: Educational Testing Service.
- Barton, Paul E., and Irwin S. Kirsch. 1990. "Workplace Competencies: The Need to Improve Literacy and Employment Readiness." Washington, DC: Office of Educational Research and Improvement, U.S. Department of Education.
- Bell, Daniel. 1976. *The Coming of Post-Industrial Society: A Venture in Social Forecasting*. New York: Basic Books.
- Berg, Ivar. 1971. *Education and Jobs: The Great Training Robbery*. Boston: Beacon Press.
- Berman, Eli, John Bound, and Zvi Griliches. 1994. "Changes in the Demand for Skilled Labor within U.S. Manufacturing: Evidence from the Annual Survey of Manufactures." *Quarterly Journal of Economics* 109: 367–397.
- Berryman, Sue E. 1993. "Learning for the Workplace." *Review of Research in Education* 19: 343–401.

- Bishop, John H. 1989. "Is the Test Score Decline Responsible for the Productivity Growth Decline?" *American Economic Review* 79: 178–197.
- Bluestone, Barry, and Bennett Harrison. 1982. *Deindustrialization of America*. New York: Basic Books.
- Bohrnstedt, George W. 1983. "Measurement." In *Handbook of Survey Research*, eds. Peter H. Rossi, James D. Wright, and Andy B. Anderson. New York: Academic Press. Pp.69–121.
- Bollen, Kenneth A. 2001. "Indicator: Methodology." In *International Encyclopedia of the Social and Behavioral Sciences*, eds. Neil J. Smelser and Paul B. Baltes. Elsevier. Pp. 7282–7287.
- Bond, Trevor G., and Christine M. Fox. 2001. *Applying the Rasch Model: Fundamental Measurement in the Human Sciences*. Mahwah, NJ: Erlbaum.
- Borghans, Lex, Francis Green, and Ken Mayhew. 2001. "Skills Measurement and Economic Analysis: An Introduction." *Oxford Economic Papers* 53: 375–384.
- Borghans, Lex, and Bas ter Weel. 2004. "Are Computer Skills the New Basic Skills? The Returns to Computer, Writing, and Math Skills in Britain." *Labour Economics* 11:85–98.
- Bosma, Hans, Michael G. Marmot, Harry Hemingway, Amanda C. Nicholson, Eric Brunner, and Stephen A. Stansfeld. 1997. "Low Job Control and Risk of Coronary Heart Disease in Whitehall II (Prospective Cohort) Study." *British Medical Journal* 314: 558–565.
- Breaugh, James A. 1998. "The Development of a New Measure of Global Work Autonomy." *Educational and Psychological Measurement* 58: 119–128.
- Breaugh, James A., and Alene S. Becker. 1987. "Further Examination of the Work Autonomy Scales: Three Studies." *Human Relations* 40: 381–400.
- Bresnahan, Timothy F., Erik Brynjolfsson, and Lorin M. Hitt. 2002. "Information Technology, Workplace Organization, and the Demand for Skilled Labor: Firm-Level Evidence." *Quarterly Journal of Economics* 117: 339–376.
- Brotheridge, Céleste M., and Raymond T. Lee. 2003. "Development and Validation of the Emotional Labour Scale." *Journal of Occupational and Organization Psychology* 76: 365–379.
- Brown, Charles. 1980. "Equalizing Differences in the Labor Market." *Quarterly Journal of Economics* 94: 113–34.
- Brown, Michael E. 1969. "Identification and Some Conditions of Organizational Involvement." *Administrative Science Quarterly* 14: 346–355.
- Buchanan, Ruth. 2002. "Lives on the Line: Low-Wage Work in the Teleservice Economy." In *Laboring Below the Line: The New Ethnography of Poverty, Low-Wage Work, and Survival in the Global Economy*, edited by Frank Munger. New York: Russell Sage Foundation. Pp.45–72.
- Bunz, Ulla. 2004. "The Computer-Email-Web (CEW) Fluency Scale: Development and Validation." *International Journal of Human-Computer Interaction* 17: 479–506.
- Cain, Pamela S., and Donald J. Treiman. 1981. "The Dictionary of Occupational Titles as a Source of Occupational Data." *American Sociological Review* 46: 253–278.

- Cappelli, Peter. 2001. "The National Employer Survey: Employer Data on Employee Practices." *Industrial Relations* 40: 635–647.
- Cappelli, Peter, and David Neumark. 2001. "Do 'High-Performance' Work Practices Improve Establishment-Level Outcomes?" *Industrial and Labor Relations Review* 54: 737–775.
- Caroli, Eve, and John van Reenen. 2001. "Skill-Biased Organizational Change? Evidence from a Panel of British and French Establishments." *Quarterly Journal of Economics* 116: 1449–1491.
- Cook, John D., S. J. Hepworth, Toby D. Wall, and Peter B. Warr. 1981. *The Experience of Work: A Compendium and Review of 249 Measures and Their Use*. New York: Academic Press.
- Cully, Mark, Stephen Woodland, Andrew O'Reilly, and Gill Dix. 1999. *Britain at Work*. New York: Routledge.
- Daly, Mary, Felix Büchel, and Greg J. Duncan. 2000. "Premiums and Penalties for Surplus and Deficit Education: Evidence from the United States and Germany." *Economics of Education Review* 19: 169–178.
- De Jonge, Jan. 1995. "Job Autonomy, Well-Being, and Health: A Study among Dutch Health Care Workers." PhD thesis. Rijksuniversiteit Limburg, Maastricht, Netherlands.
- Department of Health and Human Services. National Institute for Occupational Safety and Health. 2002. "The Changing Organization of Work and the Safety and Health of Working People: Knowledge Gaps and Research Directions." DHHS (NIOSH) Publication No. 2002-116. Cincinnati, OH.
- Department of Health and Human Services. National Institute for Occupational Safety and Health. n.d. "Stress at Work." Cincinnati, OH.
- Dickerson, Andrew, and Francis Green. 2004. "The Growth and Valuation of Computing and Other Generic Skills." *Oxford Economic Papers* 56: 371–406.
- Dolton, Peter, and Gerald Makepeace. 2004. "Computer Use and Earnings in Britain." *The Economic Journal* 114: C117–C129.
- Duncan, Greg J., and Saul D. Hoffman. 1981. "The Incidence and Wage Effects of Overeducation." *Economics of Education Review* 1: 75–86.
- Edwards, Richard. 1979. *Contested Terrain: The Transformation of the Workplace in the Twentieth Century*. New York: Basic Books.
- England, Paula. 1992. *Comparable Worth: Theories and Evidence*. New York: Aldine de Gruyter.
- England, Paula. 2005. "Emerging Theories of Carework." *Annual Review Sociology* 31: 381–399.
- European Foundation for the Improvement of Living and Working Conditions. 2006. "Work-Related Stress." Dublin, Ireland.
- Felstead, Alan, Duncan Gallie, and Francis Green. 2002. "Work Skills in Britain 1986–2001." Centre for Labour Market Studies, University of Leicester.
<http://www.skillsbase.dfes.gov.uk/downloads/WorkSkills1986-2001.doc> (accessed 2/8/03)

- Fernandez, Roberto M. 2001. "Skill-Biased Technological Change and Wage Inequality: Evidence from a Plant Retooling." *American Journal of Sociology* 107: 273–320.
- Fields, Dail L. 2002. *Taking the Measure of Work*. Thousand Oaks, CA: SAGE.
- Form, William. 1987. "On the Degradation of Skills." *Annual Review of Sociology* 13: 29–47.
- Friedman, Isaac A. 1999. "Teacher-Perceived Work Autonomy: The Concept and its Measurement." *Educational and Psychological Measurement* 59: 58–76.
- Gallie, Duncan. 1997. "Employment, Unemployment and the Quality of Life: The Employment in Europe Survey 1996." Report prepared for the European Commission.
- Gallie, Duncan, Alan Felstead, and Francis Green. 2004. "Changing Patterns of Task Discretion in Britain." *Work, Employment, and Society* 18: 243–266.
- Gerhart, Barry. 1987. "Sources of Variance in Incumbent Perceptions of Job Complexity." *Journal of Applied Psychology* 73: 154–162.
- Gerhart, Barry, Patrick M. Wright, and Gary C. McMahan. 2000. "Measurement Error in Research on the Human Resources and Firm Performance Relationship: Further Evidence and Analysis." *Personnel Psychology* 53: 855–872.
- Glick, William H., G. Douglas Jenkins, Jr., and Nina Gupta. 1986. "Method versus Substance: How Strong Are Underlying Relationships between Job Characteristics and Attitudinal Outcomes?" *Academy of Management Review* 29: 441–464.
- Glomb, Theresa M., John D. Kammeyer-Mueller, Maria Rotundo. 2004. "Emotional Labor Demands and Compensating Wage Differentials." *Journal of Applied Psychology* 89: 700–714.
- Goddard, John. 2001. "New Moon or Bad Moon Rising? Large Scale Government Administered Workplace Surveys and the Future of Canadian IR Research." *Relations Industrielles/Industrial Relations* 56: 3–33.
- Goodson, Ivor F., and J. Marshall Mangan. 1996. "Computer Literacy as Ideology." *British Journal of the Sociology of Education* 17: 65–79.
- Graham, Laurie. 1993. "Inside a Japanese Transplant: A Critical Perspective." *Work and Occupations* 20: 147–173.
- Green, Francis. 2004. "Why Has Work Effort Become More Intense?" *Industrial Relations* 43: 709–741.
- Green, Francis. 2006. *Demanding Work: The Paradox of Job Quality in the Affluent Economy*. Princeton: Princeton University Press.
- Green, Francis, and Donna James. 2003. "Assessing Skills and Autonomy: The Job Holder versus the Line Manager." *Human Resource Management Journal* 13: 63–77.
- Haahr, Jens Henrik, Hanne Shapiro, Signe Sørensen, Cathleen Stasz, Erik Frinking, Christian van't Hof, Francis Green, Ken Mayhew, and Rosa Fernandez. 2004. "Defining a Strategy for the Direct Assessment of Skills." Danish Technological Institute, Rand and Skope.

- Halaby, Charles N. 1994. "Overeducation and Skill Mismatch" *Sociology of Education* 67: 47–59.
- Handel, Michael J. 2000. "Models of Economic Organization and the New Inequality in the United States." Unpublished doctoral dissertation, Sociology Department, Harvard University.
- Handel, Michael J. 2003a. "Implications of Information Technology for Employment, Skills, and Wages: A Review of Recent Research." Arlington, VA: SRI International. Available at <http://www.sri.com/policy/csted/reports/sandt/it/>.
- Handel, Michael J. 2003b. "Skills Mismatch in the Labor Market." *Annual Review of Sociology* 29: 135–165.
- Handel, Michael J. 2004. "Implications of Information Technology for Employment, Skills, and Wages: Findings from Sectoral and Case Study Research." Arlington, VA: SRI International. Available at <http://www.sri.com/policy/csted/reports/sandt/it/>.
- Handel, Michael J. 2005a. "Trends in Perceived Job Quality, 1989–1998." *Work and Occupations* 32: 66–94.
- Handel, Michael J. 2005b. "Worker Skills and Job Requirements: Is There A Mismatch?" Washington, DC: Economic Policy Institute.
- Handel, Michael J. 2006. "The Effect of Participative Work Systems on Employee Earnings." *Research in the Sociology of Work* 16: 55–84.
- Handel, Michael J. 2007. "Computers and the Wage Structure." *Research in Labor Economics* 26: 155–196.
- Handel, Michael J., and Maury Gittleman. 2004. "Is There a Wage Payoff to Innovative Work Practices?" *Industrial Relations* 43: 67–97.
- Handel, Michael J., and David I. Levine. 2004. "The Effects of New Work Practices on Workers." *Industrial Relations* 43: 1–43.
- Harley, Bill. 2002. "Employee Responses to High Performance Work System Practices: An Analysis of the AWIRS95 Data." *Journal of Industrial Relations* 44: 418–434.
- Harrison, Bennett. 1994. *Lean and Mean: The Changing Landscape of Corporate Power in the Age of Flexibility*. New York: Basic Books.
- Harrison Bennett, and Barry Bluestone. 1988. *The Great U-Turn: Corporate Restructuring and the Polarizing of America*. New York: Basic Books.
- Hart, Sandra G., and Lowell E. Staveland. 1988. "Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research." In *Human Mental Workload*, eds. Peter A. Hancock and Najmedin Meshkati. New York: Elsevier. Pp.139–183.
- Hartmann, Heidi I., Robert E. Kraut, and Louise A. Tilly, eds. 1986–1987. *Computer Chips and Paper Clips: Technology and Women's Employment*. Washington, DC: National Academy Press.

- Harvey, Robert J. 1991. "Job Analysis." In *Handbook of Industrial and Organizational Psychology*, eds. Marvin D. Dunnette and Leaetta M. Hough. Palo Alto, CA: Consulting Psychologists Press. Pp. 71–163.
- Harvey, Robert J. 2004. "Empirical Foundations for the Things-Data-People Taxonomy of Work." Paper presented at "Things, Data, and People: Fifty years of a Seminal Theory" symposium at the annual conference of the Society for Industrial and Organizational Psychology, Chicago.
- Heinssen, Robert K., Carol R. Glass, and Luanne A. Knight. 1987. "Assessing Computer Anxiety: Development and Validation of the Computer Anxiety Rating Scale." *Computers in Human Behavior* 3: 49–59.
- Hirschhorn, Larry. 1984. *Beyond Mechanization*. Cambridge, MA: MIT Press.
- Hochschild, Arlie R. 1983. *The Managed Heart: Commercialization of Human Feeling*. Berkeley, CA: University of California Press.
- Holzer, Harry J. 1996. *What Employers Want: Job Prospects for Less Educated Workers*. New York: Russell Sage.
- Holzer, Harry J., and Michael A. Stoll. 2001. "Employers and Welfare Recipients: The Effects of Welfare Reform in the Workplace." Public Policy Institute of California, San Francisco, CA.
- Howell, David R., and Edward N. Wolff. 1991. "Trends in the Growth and Distribution of Skills in the U.S. Workplace, 1960–1985." *Industrial and Labor Relations Review* 44: 486–502.
- Hull, Frank M., and Paul D. Collins. 1987. "High-Technology Batch Production Systems: Woodward's Missing Type." *Academy of Management Journal* 30: 786–797.
- Jackall, Robert. 1988. *Moral Mazes*. New York: Oxford University Press.
- Johnston William B., and Arnold E. Packer. 1987. *Workforce 2000: Work and Workers for the 21st Century*. Indianapolis, IN: Hudson Institute.
- Kalleberg, Arne L., ed. 1986. "America at Work: National Surveys of Employees and Employers." New York: Social Science Research Council.
- Kalleberg, Arne L., and James R. Lincoln. 1988. "The Structure of Earnings Inequality in the United States and Japan." *American Journal of Sociology* 94: S121–S153.
- Kanter, Rosabeth Moss. 1977. *Men and Women of the Corporation*. New York: Basic Books.
- Karasek, Robert A. Jr. 1979. "Job Demands, Job Decision Latitude, and Mental Strain: Implications for Job Redesign." *Administrative Science Quarterly* 24: 285–308.
- Kasl, Stanislav V. 1998. "Measuring Job Stressors and Studying the Health Impact of the Work Environment: An Epidemiologic Commentary." *Journal of Occupational Health Psychology* 3: 390–401.
- Katz, Lawrence F. 2000. "Technological Change, Computerization, and the Wage Structure." In *Understanding the Digital Economy*, eds. Erik Brynjolfsson and Brian Kahin. Cambridge, MA: MIT Press. Pp.217–244.

- Katz, Lawrence F., and Kevin M. Murphy. 1992. "Changes in Relative Wages, 1963–1987: Supply and Demand Factors." *Quarterly Journal of Economics* 107: 35–78.
- Keep, Ewart, and Kenneth Mayhew. 1996. "Evaluating the Assumptions that Underlie Training Policy." In *Acquiring Skills: Market Failures, Their Symptoms, and Policy Responses*, eds. Alison L. Booth and Dennis J. Snower. Cambridge: Cambridge University Press. Pp. 305–34.
- Kelley, Jonathan, and M. D. R. Evans. 1995. "Class and Class Conflict in Six Western Nations." *American Sociological Review* 60: 157–178.
- Kirsch Irwin S., Anne Jungeblut, Lynn Jenkins, and Andrew Kolstad. 1993. *Adult Literacy in America: A First Look at the Results of the National Adult Literacy Survey*. National Center for Education Statistics. U.S. Department of Education. Washington, DC.
- Klare, George R. 1974–1975. "Assessing Readability." *Reading Research Quarterly* 10: 62–102.
- Klare, George R. 2000. "Readable Computer Documentation." *ACM Journal of Computer Documentation* 24: 148–168.
- Kohn, Melvin L., and Carmi Schooler. 1983. *Work and Personality*. Norwood, NJ: Ablex.
- Krahn Harvey, and Graham S. Lowe. 1998. *Literacy Utilization in Canadian Workplaces*. Ottawa: Statistics Canada.
- Kristensen, Tage S., Jakob B. Bjorner, Karl B. Christensen, and Vilhelm Borg. 2004. "The Distinction between Work Pace and Work Hours in the Measurement of Quantitative Demands at Work." *Work and Stress* 18: 305–322.
- Krueger, Alan B. 1993. "How Computers Have Changed the Wage Structure: Evidence from Microdata, 1984–1989." *Quarterly Journal of Economics* 108: 33–61.
- Land, Kenneth C. 1983. "Social Indicators." *Annual Review of Sociology* 9: 1–26.
- Leckie, Norm, André Léonard, Julie Turcotte, and David Wallace. 2001. "Employer and Employee Perspectives on Human Resource Practices." Ottawa: Statistics Canada.
- Leidner, Robin. 1993. *Fast Food, Fast Talk: Service Work and the Routinization of Everyday Life*. Berkeley: University of California Press.
- Levin, Henry. 1998. "Educational Performance Standards and the Economy." *Educational Research* 24: 4–10.
- Levy, Frank, and Richard J. Murnane. 1996. "With What Skills Are Computers a Complement?" *American Economic Review: Papers and Proceedings* 86: 258–62.
- Levy, Frank, Anne Beamish, Richard J. Murnane, and David Autor. 1999. "Computerization and Skills: Examples from a Car Dealership." Manuscript, Massachusetts Institute of Technology.
- Liker, Jeffrey K., Carol J. Haddad, and Jennifer Karlin. 1999. "Perspectives on Technology and Work Organization." *Annual Review of Sociology* 25: 575–596.

- Liu, Cong, Paul E. Spector, and Steve M. Jex. 2005. "The Relation of Job Control with Job Strains: A Comparison of Multiple Data Sources." *Journal of Occupational and Organizational Psychology* 78: 325–336.
- Lopata, Helen Znanieck, Kathleen Fordham Norr, Debra Barnewolt, and Cheryl Allyn Miller. 1985. "Job Complexity as Perceived by Workers and Experts." *Work and Occupations* 12: 395–415.
- Loyd, Brenda H., and Clarice Gressard. 1984. "Reliability and Factorial Validity of Computer Attitude Scales." *Educational and Psychological Measurement* 44: 501–505.
- Ludlow, Larry H. 1999. "The Structure of the Job Responsibilities Scale: A Multimethod Analysis." *Educational and Psychological Measurement* 59: 962–975.
- Manly, Donna, Cindy Bentley-Knickrehm, Lisa Flesch, and Kelly Kornacki. 1994. "Workplace Educational Skills Analysis: Training Guide Supplement." Center on Education and Work, University of Wisconsin, Madison, WI.
- Manson, Todd M., Edward L. Levine, and Michael T. Brannick. 2000. "The Construct Validity of Task Inventory Ratings: A Multitrait-Multimethod Analysis." *Human Performance* 13: 1–22.
- McGovern, Patrick, Deborah Smeaton, and Stephen Hill. 2004. "Bad Jobs in Britain: Nonstandard Employment and Job Quality." *Work and Occupations* 31: 225–249.
- McGregor, Douglas. 1957. "The Human Side of Enterprise." *Management Review* 46: 22–28.
- McIntosh, Steven, and Hilary Steedman. 2000. "Low Skills: A Problem for Europe." Final Report to DGXII of the European Commission on the NEWSKILLS Programme of Research. Centre for Economic Performance. London School of Economics and Political Science.
- McNett, Ian. 1981. "Federal Research Lab Pursues Work-Personality Connection" *APA Monitor* 12 (April): 1ff.
- Melcher, Arlyn J., Moutaz Khouja, and David E. Booth. 2002. "Toward a Production Classification System." *Business Process Management Journal* 8: 53–79.
- Meulman, Jacqueline J., Anita J. Van der Kooij, and Willem J. Heiser. 2004. "Principal Components Analysis with Nonlinear Optimal Scaling Transformations for Ordinal and Nominal Data." In *Handbook of Quantitative Methods in the Social Sciences*, edited by David W. Kaplan. Newbury Park, CA: Sage Publications. Pp. 49–70.
- Milkovich, George T., and Jerry M. Newman. 1993. *Compensation*. Homewood, IL: Irwin.
- Miller, Ann R., Donald J. Treiman, Pamela S. Cain, and Patricia A. Roos. 1980. *Work, Jobs, and Occupations: A Critical Review of the Dictionary of Occupational Titles*. Washington, DC: National Academy Press.
- Mishel, Lawrence, and Jared Bernstein. 1998. "Technology and the Wage Structure: Has Technology's Impact Accelerated Since the 1970s?" *Research in Labor Economics* 17: 305–355.
- Morrill, Calvin. 1995. *The Executive Way: Conflict Management in Corporations*. Chicago: University of Chicago Press.

- Mosenthal, Peter B. 1998. "Defining Prose Task Characteristics for Use in Computer-Adaptive Testing and Instruction." *American Educational Research Journal* 35: 269–307.
- Mosenthal, Peter B., and Irwin S. Kirsch. 1998. "A New Measure for Assessing Document Complexity: The PMOSE/IKIRSCH Document Readability Formula." *Journal of Adolescent and Adult Literacy* 41: 638–657.
- Moss, Philip I., and Chris Tilly. 2001. *Stories Employers Tell: Race, Skill, and Hiring in America*. New York: Russell Sage.
- Murnane, Richard J., and Frank Levy. 1996. *Teaching the New Basic Skills: Principles For Educating Children to Thrive in a Changing Economy*. New York: Free Press.
- Myles, John. 1988. "The Expanding Middle: Some Canadian Evidence on the Deskilling Debate." *Canadian Review of Sociology and Anthropology* 25: 335–364.
- National Research Council. Steering Committee on Research Opportunities Relating to Economic and Social Impacts of Computing and Communications. 1998. *Fostering Research on the Economic and Social Impacts of Information Technology*. Washington, DC: National Academy Press.
- Organisation for Economic Co-operation and Development. 2005. "Learning a Living: First Results of the Adult Literacy and Life Skills Survey." Ottawa and Paris: Statistics Canada and Organisation for Economic Co-operation and Development.
- Osterman, Paul. 2000. "Work Reorganization in an Era of Restructuring: Trends in Diffusion and Effects on Employee Welfare." *Industrial and Labor Relations Review* 53: 179–196.
- Paoli, Pascal. 1992. "First European Survey on the Work Environment 1991–1992." Dublin, Ireland: European Foundation for the Improvement of Living and Working Conditions.
- Parent-Thirion, Agnès, Enrique Fernández Macías, John Hurley, and Greet Vermeulen. 2007. "Fourth European Working Conditions Survey." Dublin: European Foundation for the Improvement of Living and Working Conditions.
- Payne, Jonathan. 1999. "All Things to All People: Changing Perceptions of 'Skill' among Britain's Policy Makers Since the 1950s and Their Implications." SKOPE Research Paper No. 1. Oxford University.
- Peterson, Norman G., Michael D. Mumford, Walter C. Borman, P. Richard Jeanneret, and Edwin A. Fleishman. 1999. *An Occupational Information System for the 21st Century: The Development of O*NET*. Washington, DC: American Psychological Association.
- Peterson, Norman G., Michael D. Mumford, Walter C. Borman, P. Richard Jeanneret, Edwin A. Fleishman, Kerry Y. Levin, Michael A. Campion, Melinda S. Mayfield, Frederic P. Morgeson, Kenneth Pearlman, Marilyn K. Gowing, Anita R. Lancaster, Marilyn B. Silver, and Donna M. Dye. 2001. "Understanding Work Using the Occupational Information Network (O*NET): Implications for Practice and Research." *Personnel Psychology* 54: 451–492.
- Piketty, Thomas, and Emmanuel Saez. 2003. "Income Inequality in the United States, 1913–1998." *Quarterly Journal of Economics* 118: 1–39.

- Piore, Michael J., and Charles Sabel. 1984. *The Second Industrial Divide*. New York: Basic Books.
- President's Information Technology Advisory Committee (PITAC). 1999. "Information Technology Research: Investing in Our Future." Report to the President of the United States.
- Quinn, Robert P., and Graham L. Staines. 1979. *The 1977 Quality of Employment Survey*. Ann Arbor: University of Michigan, Institute for Social Research, Survey Research Center.
- Riesman, David. 1950. *The Lonely Crowd*. New Haven: Yale University Press.
- Reich, Robert. 1991. *The Work of Nations: Preparing Ourselves for 21st-Century Capitalism*. New York: A.A. Knopf.
- Rosenbaum, James, and Amy Binder. 1997. "Do Employers Really Need More Educated Youth?" *Sociology of Education* 70: 68–85.
- Rotundo, Maria, and Paul R. Sackett. 2004. "Specific Versus General Skills and Abilities: A Job Level Examination of Relationships with Wage." *Journal of Occupational and Organizational Psychology* 77: 127–148.
- Schmenner, Roger W. 1987. *Production/Operations Management*. Chicago: Science Research Associates.
- Schultz, Kenneth S., and David J. Whitney. 2005. *Measurement Theory in Action*. Thousand Oaks, CA: Sage.
- Shaw, Kathryn. 2002. "By What Means Does Information Technology Affect Employment and Wages? The Value of IT, HRM Practices, and Knowledge Capital." In *Productivity, Inequality, and the Digital Economy*, eds. Nathalie Greenan, Yannick L'Horty, and Jacques Mairesse. Cambridge, MA: MIT Press.
- Siegel, Donald S. 1999. *Skill-Biased Technological Change: Evidence from a Firm-Level Survey*. Kalamazoo, MI: W.E. Upjohn Institute for Employment Research.
- Siegrist, Johannes, Dagmar Starke, Tarani Chandola, Isabelle Godin, Michael Marmot, Isabelle Niedhammer, and Richard Peter. 2004. "The Measurement of Effort–Reward Imbalance at Work: European Comparisons." *Social Science and Medicine* 58: 1483–1499.
- Sijtsma, Klaas, and Ivo W. Molenaar. 2002. *Introduction to Nonparametric Item Response Theory*. Thousand Oaks, CA: Sage.
- Smith, Vicki. 1997. "New Forms of Work Organization." *Annual Review of Sociology* 23: 315–339.
- Spector, Paul E., and Suzy Fox. 2003. "Reducing Subjectivity in the Assessment of the Job Environment: Development of the Factual Autonomy Scale (FAS)." *Journal of Organizational Behavior* 24: 417–432.
- Spenner, Kenneth I. 1979. "Temporal Changes in Work Content." *American Sociological Review* 44: 968–975.
- Spenner, Kenneth I. 1980. "Occupational Characteristics and Classification Systems." *Sociological Methods and Research* 9: 239–264.

- Spenner, Kenneth I. 1983. "Deciphering Prometheus: Temporal Change in the Skill Level of Work." *American Sociological Review* 48: 824–837.
- Spenner, Kenneth I. 1990. "Skill: Meanings, Methods, and Measures." *Work and Occupations* 17: 399–421.
- Spitz-Oener, Alexandra. 2006. "Technical Change, Job Tasks, and Rising Educational Demands: Looking Outside the Wage Structure." *Journal of Labor Economics* 24: 235–270.
- Statistics Canada. 2004. "Guide to the Analysis of the Workplace and Employee Survey 2002." Ottawa: Statistics Canada.
- Statistics Canada. 2005. "Measuring Adult Literacy and Life Skills: New Frameworks for Assessment." Ottawa: Statistics Canada.
- Steinberg, Ronnie J. 1990. "Social Construction of Skill: Gender, Power, and Comparable Worth." *Work and Occupations* 17: 449–482.
- Steinberg, Ronnie J., and Deborah M. Figart. 1999. "Emotional Demands at Work: A Job Content Analysis." *ANNALS of the American Academy of Political and Social Science* 561: 177–191.
- Stenner, A. Jackson, and Mark H. Stone. 2004. "Does the Reader Comprehend the Text Because the Reader Is Able or Because the Text Is Easy?" Paper presented at the International Reading Association, Reno-Tahoe, Nevada, May, 2004. Durham, NC: Metametrics.
- Sticht, Thomas G. 1975. *Reading for Working: A Functional Literacy Anthology*. Human Resources Research Organization, Alexandria, VA.
- Stone, Katherine. 1975. "The Origin of Job Structures in the Steel Industry." In *Labor Market Segmentation*, eds. Richard Edwards, Michael Reich, and David Gordon. Lexington, MA: D.C. Heath. Pp.27–84.
- Sum, Andrew M. 1999. *Literacy in the Labor Force*. Washington, DC: National Center for Educational Statistics, U.S. Department of Education.
- Swidler, Ann. 1986. "Culture in Action." *American Sociological Review* 51: 273–286.
- Tam, Tony. 1997. "Sex Segregation and Occupational Gender Inequality in the United States: Devaluation or Specialized Training?" *American Journal of Sociology* 102: 1652–1692.
- Tilly, Chris, and Charles Tilly. 1997. *Working Under Capitalism*. Boulder, CO: Westview.
- U.S. Department of Labor. Employment and Training Administration. 1991. *The Revised Handbook for Analyzing Jobs*. Washington, DC: GPO.
- U.S. Department of Labor. Employment and Training Administration. 1993. "The New DOT: A Database of Occupational Titles for the Twenty-First Century." Final Report of the Advisory Panel for the *Dictionary of Occupational Titles* (APDOT).
- U.S. Department of Labor. Employment and Training Administration. 2005. "O*NET Data Collection Program." Office of Management and Budget Clearance Package Supporting Statement and Data Collection Instruments. Washington, DC.

- U.S. Department of Labor. Secretary's Commission on Achieving Necessary Skills. 1991. *What Work Requires of Schools: A SCANS Report for America 2000*. Washington, DC: U.S. Department of Labor.
- U.S. National Commission on Excellence in Education. 1983. *A Nation at Risk: The Imperative for Educational Reform*. Washington, DC: Government Printing Office.
- Vallas, Steven Peter. 1990. "The Concept of Skill: A Critical Review." *Work and Occupations* 17: 379–398.
- Vallas, Steven Peter. 1999. "Rethinking Post-Fordism: The Meaning of Workplace Flexibility." *Sociological Theory* 17: 68–101.
- Vallas, Steven Peter. 2003. "Why Teamwork Fails: Obstacles to Workplace Change in Four Manufacturing Plants." *American Sociological Review* 68: 223–250.
- van de Ven, Andrew, and Diane L. Ferry. 1980. *Measuring and Assessing Organizations*. New York: Wiley.
- van Schuur, Wijbrandt H. 2003. "Mokken Scale Analysis: Between the Guttman Scale and Parametric Item Response Theory." *Political Analysis* 11: 139–163.
- Wall, Toby D., Paul R. Jackson, and Sean Mullarkey. 1995. "Further Evidence on Some New Measures of Job Control, Cognitive Demand, and Production Responsibility." *Journal of Organizational Behavior* 16: 431–455.
- Walton, Richard E. 1974. "Alienation and Innovation in the Workplace." In *Work and the Quality of Life: Resource Papers for Work in America*, edited by James O'Toole. Cambridge, MA: MIT Press. Pp. 227–245.
- Wharton, Amy. 1999. "The Psychosocial Consequences of Emotional Labor." *The ANNALS of the American Academy of Political and Social Science* 561: 158–176.
- White, Michael, et al. 2004. "Changing Employment Relationships, Employment Contracts and the Future of Work, 1999–2002." (4741userguide2.pdf). Colchester, Essex: UK Data Archive. Study Number 4641. <http://www.data-archive.ac.uk/findingData/snDescription.asp?sn=4641> (accessed 2/27/07).
- Wilson, Mark, Robert J. Harvey, and Barry A. Macy. 1990. "Repeating Items to Estimate the Test-Retest Reliability of Task Inventory Ratings." *Journal of Applied Psychology* 75: 158–163.
- Wilson, William Julius. 1987. *The Truly Disadvantaged*. Chicago: University of Chicago Press.
- Wilson, William Julius. 1996. *When Work Disappears: The World of the New Urban Poor*. New York: Knopf.
- Wise, Laress, Wei Jing Chia, and Lawrence M. Rudner. 1990. "Identifying Necessary Job Skills: A Review of Previous Approaches." Pelavin Associates, Washington, DC.
- Woodward, Joan. 1965. *Industrial Organization: Theory and Practice*. New York : Oxford University Press.

Wright, Erik Olin. 1985. *Classes*. London: Verso.

Wright, Erik Olin. 1997. "Rethinking, Once Again, the Concept of Class Structure." In *Reworking Class*, edited by John R. Hall. Ithaca, NY: Cornell University Press. Pp.41–72.

Wright, Erik Olin, David Hachen, Cynthia Costello, and Joey Sprague. 1982. "The American Class Structure." *American Sociological Review* 47: 709–726.

Zuboff, Shoshana. 1988. *In the Age of the Smart Machine: The Future of Work and Power*. New York: Basic Books.