Research Monograph: Development of the Common-Metric Questionnaire(CMQ)

Robert J. Harvey
Personnel Systems & Technologies Corporation

Abstract

The Common-Metric Questionnaire (CMQ), an incumbent-completable job analysis instrument targeted at both exempt and nonexempt jobs, was developed to address several limitations facing past worker-oriented instruments, including: (a) reading levels that are so high that they preclude incumbent completion; (b) items and rating scales that are so behaviorally abstract that it is difficult to collect accurate and verifiable data; and (c) deficiencies in content coverage, especially for managerial jobs. The CMQ's items and rating scales were written to be both more behaviorally specific than typical the worker-oriented instrument (2,077 pieces of item-level data are collected) and easier to understand (an eighth-grade reading level was used). The CMQ's "matrix" structure allows efficient data collection: Approximately two-thirds of the 4,552 positions (representing over 900 jobs) analyzed in the field test were rated in 2 hours or less. Over 1,200 CMQ ratings are used in a scoring system that describes jobs in terms of 80 work dimensions (median scale coefficient alpha = .86). CMQ work dimension scores are useful for functions that require an abstract view of work (e.g., compensation, job classification, synthetic validity); for addressing functions that require greater specificity (e.g., job descriptions, performance appraisals), CMQ item-level data can be used. The CMQ dimensions are highly predictive of external criteria: in the CMQ field-test, policy-capturing Rs predicting pay rates were in the .80's and .90's, and hit rates predicting dichotomous FLSA exempt status were in the 80-90% range.

Background

The goal of "worker-oriented" job analysis is to be able to analyze -- and meaningfully compare -- even highly task-dissimilar jobs; this is accomplished by describing each job using a common profile, or metric, of work activities (e.g., see Harvey, 1991a; McCormick, 1976). Worker- oriented instruments describe jobs in terms of their general work behaviors (GWBs); these are characteristics of the work or work environment that tend to be more behaviorally abstract than the items produced in a task- based job analysis. By using this more behaviorally abstract metric, underlying similarities can be identified between jobs that might otherwise appear to be highly dissimilar if analyzed at the task level.

The worker-oriented approach to job analysis has been used for many years (e.g., Cunningham, 1964; Palmer &

McCormick, 1961), and many standardized instruments that implement this philosophical approach to work measurement have been developed (e.g., the Job Element Inventory, or JEI, of Cornelius & Hakel, 1978; the General Work Inventory, or GWI, of Cunningham & Ballentine, 1982; the Occupation Analysis Inventory, or OAI, of Cunningham, Boese, Neeb, & Pass, 1983; the Executive Checklist, or EXCEL, of Lozada-Larsen, 1988; the Position Analysis Questionnaire, or PAQ, of McCormick, Jeanneret, & Mecham, 1972; the Professional and Managerial Position Questionnaire, or PMPQ, of Mitchell & McCormick, 1978). These instruments have been used to address a variety of personnel needs, including grouping jobs for classification purposes, conducting synthetic (or job-component) validation studies (e.g., McCormick, 1959, and setting compensation rates.

However, existing worker-oriented instruments suffer from a number of potentially serious limitations, both in terms of their psychometric characteristics and their usefulness in solving personnel problems. Regarding psychometric limitations, these have been discussed in detail elsewhere (e.g., Harvey, 1991a, pp. 81-98), and the main areas of concern will be summarized in the next section. Regarding usability limitations, the fact that most existing instruments rate jobs using a small number of highly abstract items renders them largely unusable for a number of personnel functions. For example, developing job descriptions and employee performance appraisal forms requires relatively detailed data, and one seldom sees worker-oriented instruments used for such purposes.

The Common-Metric Questionnaire (CMQ; Harvey, 1990; 1991b) was developed to address a number of the limitations that, in varying degrees, face existing worker- oriented instruments. In short, I attempted to develop an instrument that would (a) be easier for raters to understand and complete, (b) describe more specific, observable aspects of jobs (e.g., to facilitate the verification of the ratings, and allow for the development of job descriptions and appraisal forms), and (c) describe both nonexempt and exempt (e.g., managerial, executive, and professional) jobs using a common profile of work activities.

Of course, it must be stressed that the CMQ was not designed to provide the ultimate solution to every job analysis problem. Although worker-oriented instruments are useful in addressing a wide variety of functions, there will always be situations in which highly specific task- or element-level data are required (e.g., developing detailed training programs). Even in these situations, however, the CMQ can still be valuable by providing an initial framework from which the detailed task-level analysis can begin, thereby producing potentially important savings in labor relative to conducting a task analysis "from scratch."

The sections below summarize the limitations of existing job analysis instruments, the CMQ design goals, the process of developing the CMQ, and the results of the CMQ field-test.

Limitations of Existing Standardized Job Analysis Instruments

Over 30 years ago (e.g., Palmer & McCormick, 1961; McCormick, Cunningham, & Gordon, 1967), the efforts of Ernest McCormick and his associates to develop worker-oriented job analysis instruments effectively revolutionized the theory and practice of job analysis. McCormick proposed developing standardized instruments that would be capable of describing and comparing even highly task-dissimilar jobs using a common profile of more abstract items. The potential advantages of this approach were twofold: (a) achieving significant cost savings via standardization of the instrument (i.e., the expense of developing a job analysis questionnaire would be avoided); and (b) allowing meaningful, quantitative comparisons to be made between jobs that differ at the task level.

Unfortunately, the worker-oriented instruments that have been developed over the past 30 years suffer from various design- and implementation-based limitations that may seriously restrict their ability to achieve the above goals of data-collection economy and cross-job comparability. Two general classes of limitations can be identified: (a) problems with respect to providing a common metric of GWBs on which to describe jobs; and (b) problems with the rating scales used to rate the GWB items. The net result of these limitations is that many worker-oriented instruments are difficult or expensive to use, and their ratings may fall short of the goal of providing a common metric that identifies meaningful level-based differences between jobs.

Item-Based Limitations

With respect to limitations based on the items used to rate jobs, a number of specific issues can be identified. First, content coverage may be lacking: general-purpose instruments (e.g., JEI, OAI, PAQ) designed to analyze all types of jobs tend to lack coverage of the work activities that characterize executive, professional, and managerial jobs. In contrast, those instruments that have been specifically targeted at higher-level occupations (e.g., PMPQ, EXCEL) tend to lack coverage of the breadth of activities performed in nonmanagerial jobs.

McCormick acknowledged the coverage limitations facing the PAQ, observing that it has "difficulties with higher level positions" due to the "level of generality" of its items (Mitchell & McCormick, 1979, p. 4). Mitchell and McCormick further speculated that the format of the rating scales used in the PAQ "severely restricts the range of possible responses and may not be sufficiently discriminating at the upper levels" (p. 5).

However, instead of simply expanding the PAQ's item pool to include items to span these missing managerial dimensions, McCormick opted to construct an entirely new instrument targeted solely at higher-level jobs. The result of this effort was the PMPQ (Mitchell & McCormick, 1979); other researchers took a similar approach (e.g., the Position Description Questionnaire, or PDQ, of Page & Gomez, 1979).

Unfortunately, these managerially targeted instruments do not solve the original problems: (a) they are not suitable for describing nonmanagerial jobs, (b) they do nothing to remedy the content deficiencies of the general purpose instruments, and, perhaps of greatest importance, (c) they do not allow managerial and nonmanagerial jobs to be compared on a common metric. The fact that lower- and upper-level jobs are described using a different metric (i.e., different job analysis items and rating scales) makes it impossible to compute quantitative comparisons between entire classes of jobs, thereby removing the primary benefit of worker-oriented analysis!

Second, many worker-oriented instruments contain items that are written at an extremely high level of behavioral abstraction. I have argued (e.g., Harvey, 1991a; Harvey, Wilson, & Blunt, 1993) that these abstract, "holistic" items are -- by their very nature -- very difficult to rate accurately, and they additionally pose serious problems when one attempts to verify the accuracy of the individual job analysis item ratings. For example, the PAQ contains single-item scales to rate such complex activities as "criticality of position," "job structure," and "general responsibility;" the PMPQ rates items including "oral communications," "judgments involving people," "using equipment and devices," and "supervising/directing."

The practice of using such behaviorally abstract items in worker-oriented instruments does nothing to facilitate the rater's information processing task. When raters attempt to make accurate ratings, the cognitive task facing them when judging highly abstract, unobservable items is daunting. Conversely, in situations in which pressures to distort the job analysis ratings are present (e.g., when the job analysis is being conducted to set compensation rates), it is self-evident that holistic items are much more susceptible to faking or distortion -- and much less able to allow detection of dissimulation if it occurs -- than ratings of observable job characteristics.

Third, the wording of the items, rating scales, and instructions seen in many worker-oriented instruments is such that the reading level required to administer or complete the instrument is quite high. For example, Ash and Edgell (1975) estimated the PAQ's reading level to be at the post-college-graduate level. The consequences of high reading levels are that (a) job incumbents or other subject matter experts (SMEs) typically cannot serve as raters; (b) expensive job analysts or consultants are required to make the actual job analysis item ratings, contributing significantly to the overall cost of a job analysis project; and (c) the results of the job analysis are of limited use for functions that demand more specific information (e.g., developing job descriptions and performance appraisal forms). In effect, the term "worker-oriented" has become synonymous with "behaviorally nebulous" among many I/O practitioners.

I would argue that although it is true that most existing worker-oriented instruments are of very little use for personnel functions like developing job descriptions and employee appraisal forms, it is not necessarily the case that worker-oriented instruments cannot address such functions. In short, there is nothing in the basic concept of a "worker-oriented" approach to job analysis that requires either the items or item rating scales to be behaviorally vague and nebulous, or that requires raters to make speculative, difficult-to-justify, holistic judgments.

There are effectively only two essential characteristics of a worker-oriented job analysis questionnaire: (a) it must comprehensively describe work activities or characteristics using a common set of items that are written at a more behaviorally abstract level than is seen in the typical task statement; and (b) it must include rating scales that describe each item on a metric that retains a constant meaning across all jobs. The former is necessary to allow the instrument to identify cross-job similarities that may be masked by task differences; the latter is required in order to allow meaningful level-based differences and similarities across jobs to be quantified.

Given the near impossibility of validating holistic ratings of unobservable job characteristics, there is ample reason to break with the past tradition of using highly abstract items and rating scales in worker-oriented instruments. As Cunningham's OAI demonstrated, there is plenty of middle-ground between the highly abstract items seen in older worker-oriented instruments on the one hand, and the very high technological specificity present in task- and element-level items on the other. Accordingly, the CMQ was designed to contain items that are more behaviorally abstract than tasks, but less behaviorally abstract than the items seen in many worker-oriented instruments.

Rating Scale Limitations

With respect to the second category of limitations -- the methods used to rate the worker-oriented items -- several additional issues can be raised. First, some rating scales and/or rater instructions produce ratings that are within-job relative in nature (see Harvey, 1991a, pp. 97-98); such rating scales effectively prevent one from making meaningful level-based cross-job comparisons.

For example the JEI uses a relative-time-spent scale in which each item is judged relative to the time spent on the "average" task performed on the job; as a consequence, a '3' rating for one job may denote a totally different actual amount of time than the identical '3' rating given to a second job. Likewise, some instruments make use of "relative importance" scales that produce effectively ipsative ratings: i.e., the notion of importance is specific to the range of activities performed on each job, and every job will have at least some items rated high on importance.

In such cases, even if the same numerical rating is given to the same activity for two different jobs, it will often be the case that the absolute importance of the activity (e.g., expressed in terms of the consequences of improper task performance with respect to monetary, life, or property loss) may be very different. For example, an electronics assembler for a video-game company and a pilot of a commercial jetliner may both rate "activities involving eye-hand coordination" as being a "5" (Extreme importance) for their jobs. Despite the fact that the two jobs received identical numerical ratings on this item, it is obvious that the impact of ineffective performance is quite different: in the former case, an assembly error might cause a $5 video game to fail a quality control inspection; in the latter case, a multi-million dollar aircraft and hundreds of lives could be lost as a consequence of pilot error when operating the controls of the aircraft.

The fundamental issue is whether the numerical rating values on a job analysis scale "mean the same thing" with respect to the amount or level of each GWB item that is present in each job. Clearly, in this example, activities involving eye-hand coordination are not equally "important" to both jobs, yet the numerical item ratings given using the relative importance scale would indicate that they are identical. Although both jobs involve eye-hand coordination, the absolute level of importance is fundamentally different; this fact must be captured by the rating scale(s) used to describe each worker-oriented item.

Using relativistic rating scales like the one illustrated above for the purpose of making cross-job comparisons can lead to totally misleading conclusions (e.g., that electronics assemblers and airline pilots are equal with respect to the importance of eye-hand coordination). Unfortunately, within-job relativistic scales like this are commonly used in worker-oriented instruments for the purpose of making cross-job comparisons.

Second, many instruments employ behaviorally vague, imprecise rating scales and anchors to rate the GWB items; such scales pose potentially serious limitations on the degree to which the job analysis ratings can be made accurately and verifiably. For example, many PAQ items are rated on an "Importance to This Job" scale, which uses anchors that are not defined in easily verifiable terms (e.g., 3 = Average, 4 = High, 5 = Extreme importance). Unfortunately, given the lack of any reference to observable job behaviors in these scale anchors, the task of explaining why a given work activity was rated, say, Average instead of High, or High instead of Extreme importance, is potentially quite difficult and subjective. This problem is greatly exacerbated if behaviorally vague rating scales are used to rate behaviorally vague items.

Some worker-oriented instruments add an additional level of difficulty to the problem of explaining or justifying their item ratings by virtue of the fact that they do not provide a text anchor for each numeric rating point. For example, many PMPQ scales only anchor every other rating point; likewise, the PAQ allows unanchored "half point" ratings to be made between the anchored whole- numbered points (e.g., a 3.5 rating of importance would denote an unspecified level somewhere between Average and High). Because no standardized text definition corresponds to the numeric item rating, the task of explaining an unanchored job analysis rating will unavoidably be much more challenging than the task of documenting an anchored rating, regardless of the degree of behavioral ambiguity present in that anchor.

The lack of verifiability caused by using behaviorally vague or unanchored scales creates a number of other problems. Because important personnel outcomes (e.g., compensation rates) often ride on the results of the job analysis, it is not surprising to find that the employees whose jobs are being analyzed demonstrate a great deal of interest in the job analysis ratings. Unfortunately, to the degree that the ratings are defined in vague, non-behavioral terms, it is increasingly difficult to explain or defend these ratings. Indeed, the job analyst may be able to cite little beyond his or her "expert" credentials as justification for why a given job analysis item was rated the way it was. To the extent that it is deemed important for employees to have confidence in the validity of the job analysis ratings, this lack of behavioral specificity -- and thus, verifiability -- may represent a very serious problem.

Finally, the use of behaviorally vague job analysis items and/or scales may cause problems if the organization's personnel practices are challenged in court. The need for job analysis ratings to be accurate and independently verifiable has been made all the more important with the advent of legislative initiatives (e.g., the Americans with Disabilities Act of 1990, or ADA; the Civil Rights Act, or CRA, of 1991) that hold employers responsible for demonstrating the job-relatedness of their personnel decisions (in some cases allowing large punitive damages to be awarded for violations). There is good reason to expect that the courts will be taking a much closer look at the job analysis process in general, and the verifiability of individual job analysis item ratings in particular.

To summarize, existing job analysis instruments suffer from a variety of potentially serious limitations with respect to content coverage, verifiability, cross-job meaningfulness, high reading levels, and similar factors that decrease their usability and defensibility. The next section describes how the CMQ's design goals were developed in response to these limitations.

Design Goals for the CMQ

The general design goals of the CMQ were that it should (a) allow job incumbents and other non-experts to describe their own jobs and verify the ratings of other jobs; (b) have higher potential accuracy and verifiability than most existing worker-oriented instruments; and (c) produce ratings that describe exempt and nonexempt jobs on a common metric. Simply addressing the psychometric limitations of past instruments would not have been sufficient if the new instrument were so complicated that it could only be used by specialized job analysts. Conversely, making the new instrument easier to use would not have been sufficient if the ratings continued to be psychometrically questionable or difficult to verify.

The strategy taken to achieve these objectives involved using a higher level of behavioral specificity in the CMQ's items and rating scales than has been seen in most existing worker-oriented instruments. Most instruments are composed of a relatively small number of behaviorally abstract items (e.g., the JEI has 153 items, the PAQ has 194, and the PMPQ has 98). In contrast, Cunningham's OAI contains over 600 items, many of which are significantly more behaviorally specific than those seen in other instruments. OAI research (e.g., Cunningham et al., 1983) has indicated that a higher level of behavioral specificity can be achieved without sacrificing the GWB nature of the items (and therefore, retaining the ability to meaningfully compare task-dissimilar jobs).

Accordingly, it was decided that the items to be used in the CMQ would continue the trend begun by the OAI toward higher behavioral specificity, while at the same time improving on the limitations of past items (e.g., complicated wording, high reading level). Behaviorally nebulous, "holistic" items were to be avoided. The main benefit of using more behaviorally specific items is their higher intrinsic verifiability: that is, if the items are sufficiently specific, anyone familiar with the duties of the job should be able to (a) accurately rate whether each activity applies or not; (b) accurately rate the degree to which each applies; and (c) independently verify the accuracy of job analysis ratings made by other raters.

With respect to the rating scales used in the CMQ, behaviorally explicit, non-relativistic anchors were desired for all scales that were to be used to make cross-job comparisons. Although some within-job relativistic scales were included in the CMQ in order to rate the "criticality" or "essentially" of a job's work activities in the context of each organization (e.g., for ADA-compliance purposes), in no case would such ratings be used to make cross-job comparisons or use to compute the job dimension scores. Thus, at least one cross-job-relative scale was desired for each CMQ item (e.g., absolute frequency).

The ability for the CMQ to be understood and completed by untrained job incumbents or supervisors was a critical design goal; this was important for both economic (i.e., avoiding the need for expensive external job analysts) as well as "process" reasons (i.e., so that employees would be more involved in the job analysis process, and therefore more likely to accept the results of the job analysis). This necessitated using item and rating scale wording that was significantly simpler -- and more behaviorally explicit -- than that used in many existing instruments.

My experiences with the incumbent-completable JEI (e.g., Harvey, Friedman, Hakel, & Cornelius, 1989) were sufficient to convince me of the viability and desirability of designing the instrument to allow job incumbents to rate their own jobs. Accordingly, an eighth-grade reading level for the individual words used in the CMQ was set as the goal. Although the CMQ does not require incumbent completion, the capacity for incumbents and supervisors to be able to complete the instrument -- and to verify the validity of item ratings made by others -- was critical.

Finally, it was deemed necessary for the CMQ to address the content-validity problems facing many general-purpose job analysis instruments (i.e., a lack of adequate coverage of the activities that characterize executive, managerial, and professional jobs). Accordingly, an integrative listing of the work dimensions measured by previous instruments was developed, paying special attention to the dimensions that were seen only in managerially targeted instruments. This listing of work activity constructs was used to define the CMQ's content domain prior to the actual item-writing process.

Of course, achieving some of the above goals might tend to work against achieving other goals: for example, increasing the behavioral specificity of the items might also increase the number of items in the instrument, which might in turn make incumbent completion more difficult. The approach taken to this tradeoff was to use a "matrix" format for the CMQ in which the number of items (i.e., rows) was kept to a reasonable level; increased behavioral specificity was achieved by increasing the number of ratings (i.e., columns) possible for each applicable row item. Only if a particular row characteristic applies to the job would the column ratings be required, thereby minimizing the administrative demands of the instrument. As a practical matter, a goal of completing the CMQ in a reasonable amount of time (e.g., 2-3 hours or less) was set.

Developing the CMQ Dimension Targets and Item Pools

The first step in developing the items for the CMQ was to define the underlying work dimensions to be measured by the instrument. To accomplish this goal, a review was conducted of the dimensions produced by past general-purpose and occupationally targeted instruments, and an integrative dimensional framework was produced to guide the item-writing process.

Regarding the dimensions of nonmanagerial work, the PAQ has effectively constituted the standard general-purpose instrument to date. The items of the PAQ have been factored to produce several different dimensional systems: (a) the System I PAQ factor analyses (Jeanneret, 1969; McCormick et al., 1972), which produced a 5-factor solution at the overall level; (b) the System II analyses (McCormick, Jeanneret, & Mecham, 1972), which produced 13 factors; and (c) a reanalysis of the System I PAQ data (Harvey, 1987), which produced a 19-factor solution. Harvey (1991a, pp. 151-152) summarized these various factor solutions.

When identifying the general-purpose dimensions to be targeted in the CMQ, factor analytic results from the JEI (Harvey et al., 1988; Harvey, 1989) were also considered. The JEI was developed by Cornelius and Hakel (1978) to be a self-administered questionnaire; in order to make the JEI incumbent-completable, Cornelius and Hakel (a) dramatically reduced the reading level of the instrument's items relative to those seen in the PAQ, (b) used only a single response scale, and (c) shortened the length of the instrument (153 v. 194 items). To simplify matters, it was decided that only the "overall" factors (i.e., those obtained by factoring the entire pools of items) would be used to define the general-purpose work behavior constructs to be measured in the CMQ. Table 1 lists the dimensions that have been extracted from the PAQ and JEI.

Table 1. Summary of General-Purpose Work Dimensions

```
Decision making, communication, general responsibility
Operate/adjust/tend machines/tools/equip
Direct/first-line supervisory activities
Clerical/office equipment usage
External contacts
Personally demanding situations
Unpleasant environment (temperature, pollution, outdoor)
Risks & hazards
Planning & scheduling
Internal contacts
Graphic/measurement/quantitative/technical duties
Precise/repetitive activities (inverse of autonomous work)
Operate machines/tools requiring vigilance
Gross body activities (stationary, moving, balance, agility)
Exchanging information
Misc. context (optional/required apparel, nontypical schedule)
Environmental sensing/judging
Personal service duties
Regular work schedule
```

The work dimensions listed in Table 1 are consistent with the view that the general-purpose instruments from which they were derived allow relatively little depth with respect to measuring the activities that predominate in managerial occupations. The only dimension directly relevant to these functions is the "Decision-Communication-Social Responsibility" one, which is unfortunately quite general.

This lack of multiple dimensions on which higher-level jobs could be described was addressed by reviewing the dimensions produced by managerial instruments; Table 2 summarizes the dimensions uncovered via factor analysis of such instruments. More detailed reviews of the work dimensions measured by managerially targeted instruments are available elsewhere (e.g., Lozada-Larsen, 1986, 1988; Mitchell & McCormick, 1979; Page & Gomez, 1979).

Table 2. Summary of Managerial Work Dimensions

```
Internal Contacts:  Boundary Management
Internal Contacts: Internal Relations
Internal Contacts: Consulting, Communication, Information
```

Exchange
Internal Contacts: Coordination
Internal Contacts: Level/Type of Contacts
Human Resource Responsibility: Supervision
Human Resource Responsibility: Administration
Human Resource Responsibility: Conflict Resolution
Human Resource Responsibility: Instruction
Human Resource Responsibility: Delegation
Technical Activities: R&D
Production or Service Activities
Technical Support/Consulting
Emergency/Crisis Control, Safety
Resource and Asset Responsibility: Planning & Scheduling
Resource and Asset Responsibility: Allocation
Resource and Asset Responsibility: Management
Resource and Asset Responsibility: Administration
Planning/Decision Making: Current Operations
Planning/Decision Making: Future Operations
Long Range or Strategic Planning
Production or Performance Goal Setting
External Relations: Formulation and Administration of Legal
External Relations: Unions & Special Interests
External Relations: Government, Legal, Regulatory
External Relations: The Public
Training, Education, Experience
Other Language Use
Personal or Self-Development

It is interesting to note from Table 2 that many of the work dimensions uncovered from managerially targeted instruments deal almost exclusively with the general managerial factor seen in Table 1 -- that is, managerial decisions, interpersonal responsibility, and general responsibility. The work dimensions described in Tables 1-2 were integrated to form dimension-category targets for the CMQ; these categories are listed in Table 3.

When forming these general work-dimension targets, the conceptual model of work developed by Sidney Fine (e.g., Fine, 1989) proved helpful: that is, each general category of work behavior identified from prior factor analyses of worker-oriented instruments was grouped into one of Fine's categories of Data (Decision Making/Information Processing, using the labels contained in Table 3), People (Interpersonal), and Things (Physical, Mechanical, and Contextual). Prior research (e.g., Harvey, 1987) that studied the second-order factors obtained by factoring the work dimensions measured by the PAQ indicated that at this higher level of abstraction, the dimensions clustered in a fashion very similar to Fine's theory of work. Thus, to provide the most general view of the dimensionality of work, the CMQ used this conceptual structure.

Table 3. Summary of CMQ Dimension Categories

<u>Interpersonal Dimensions</u>

Kinds of supervision received
Kinds of employees supervised, closeness of supervision given
Internal contacts: kinds of workers contacted, types/purposes of contacts required
External contacts: kinds of workers contacted, types/purposes of contacts required

<u>Decision Making/Information Processing Activities</u>

Required areas of job knowledge

Language usage: kinds of languages, kinds of activities performed on each
Using senses to assess the work environment
Managerial decisions: Financial, human resources, operations/production, strategic planning
Impact/scope of managerial decisions

Physical, Mechanical, and Contextual Aspects of Work

Required physical activities
Machine/equipment usage: Kinds of machines used, kinds of activities with each
Impact of tool/machine use
Working environment conditions
Consequences/hazards of working environment conditions
Demanding/stressful job situations
Kinds/sources of performance feedback
Task/skill variety, autonomy, identity, dependence
Kinds of reward systems, work scheduling

When developing the items used in the CMQ, pools of items were developed for each of the dimension categories to be measured (listed in Tables 1-3). Once items for each dimension were written, they were pilot tested to identify potential problems (e.g., ambiguity) at an early stage; conformity of each item to the eighth-grade reading level goal was also assessed at this time.

To keep the overall length of the instrument to a minimum, the CMQ uses a "matrix" structure in which the various item content "rows" were crossed with a series of rating scale "columns" to describe the involvement of each row item in each job; the content of the row ratings varied depending on the kind of matrix being considered. For example, for the internal and external contacts matrix, the rows represented the kinds of people contacted, and the columns described the kinds of activities performed with each. In contrast, for the meetings attended/chaired matrix, the rows represented the purpose for the meeting, and the columns defined the kinds of employees who attended the meetings.

In addition to the dichotomous apply/does-not-apply (DNA) ratings obtained where the row and column elements intersect, most row items in the CMQ are also rated using a frequency scale anchored in concrete terms (e.g., the activity is performed every few hours to daily). Both the apply/DNA ratings and the multipoint frequency ratings define a metric that is comparable across all jobs rated; these are the items that are used to form the work dimension scores of the CMQ.

Finally, to provide the kind of within-job relative ratings that are useful in defining the essential functions of jobs (e.g., for ADA-compliance purposes), the CMQ row elements are also rated in terms of their importance with respect to performing the main mission of the job. Because these scales define the essentiality or criticality of each item relative to the objectives of each job considered on its own, the CMQ essentiality ratings should not be used when making cross-job comparisons (i.e., they were included in order to allow the activities performed on each job, considered individually, to be ranked in terms of their relative importance to meeting the goals of that organization).

The CMQ Field-Test

A nationwide sample of positions was collected during the CMQ field test, which was conducted during 1989-1991. In all, 4,552 positions were rated using the field-test, or "standardization" edition of the CMQ; these positions were classified into 904 separate occupational titles through analysis of the reported DOT codes and job titles.

The conditions under which the CMQ profiles were collected varied across organization: some organizations subjected each incumbent-completed CMQ booklet to a detailed review by supervisory personnel to ensure accuracy, one organization collected CMQs using a "job analyst" format (i.e., a single personnel official who was familiar with all jobs analyzed was used to collect the data and make the CMQ ratings), and the remainder used incumbent completion and an unknown degree of subsequent supervisory review of the booklets. Tables 4-18 summarize the characteristics of the field test sample

Table 4. Breakdown of sex of position incumbents in CMQ field test sample.

	Frequency	Percent
Male	2765	62.0
Female	1692	38.0

Table 5. Breakdown of whether incumbent has left and later rejoined the organization

	Frequency	Percent
No	4228	92.9
Yes	324	7.1

Table 6. Breakdown of seasonal work

	Frequency	Percent
No	3813	83.8
Yes	739	16.2

Table 7. Breakdown of reported ethnic group.

Group	Frequency	Percent
Black	549	12.2
White	3731	82.8
Hisp.	108	2.4
Asian	76	1.7
Native	16	0.4
Other	25	0.6

Table 8. Breakdown of FLSA exempt status

	Frequency	Percent
No	2665	58.5
Yes	1887	41.5

Table 9. Breakdown of percent of travel.

	Frequency	Percent
None	2434	53.5
1-25%	1616	35.5
26-50%	254	5.6
51-75%	122	2.7
76-up	126	2.8

Table 10. Breakdown of whether job requires overnight travel

	Frequency	Percent
No	3447	75.7
Yes	1105	24.3

Table 11. Breakdown of whether uniforms required.

	Frequency	Percent
No	3784	83.1
Yes	768	16.9

Table 12. Breakdown of whether protective gear is required.

	Frequency	Percent
No	3385	74.4
Yes	1167	25.6

Table 13. Breakdown of whether suits/office attire are worn.

	Frequency	Percent
No	1649	36.2
Yes	2903	63.8

&>

		Frequency	Percent
None	1	3220	70.7
1	2	1154	25.4
>1 in 1	3	101	2.2
2 areas	4	77	1.7

Table 15. Breakdown of job level/family.

	Frequency	Percent
Executive	9	0.5
Managerial	160	9.0
Supervisory	150	8.4
Prof/Technical	1100	61.6
Sales	3	0.2
Clerical	195	10.9
Farming	11	0.6
Precision	40	2.2
Labor	117	6.6

Table 16. Breakdown of difficulty ratings.

	Frequency	Percent
Very easy	561	12.5
Mod. easy	1938	43.1
Average dif.	1432	31.8
Mod. diffic.	485	10.8
Very diffic.	85	1.9

Table 17. Breakdown of time to complete the CMQ.

	Frequency	Percent
Under hour	374	8.3
1-2 hours	2580	57.1
2-3 hours	1272	28.2
3-4 hours	224	5.0
Over 4 hours	66	1.5

Table 18. Breakdown of single-digit occupational category.

1-Digit	Frequency	Percent
0.Prof/tech	773	17.7
1. " "	993	22.8
2.Clerical	1121	25.7
3.Service	796	18.2
4.Farming/fish	15	0.3
5.Processing	10	0.2
6.Machine trade	143	3.3
7.Bench work	84	1.9
8.Structural	305	7.0
9.Misc.	121	2.8

Sampling targets were not set for the demographic variables (e.g., sex, race, FLSA exempt); instead, we simply wanted as much representation of each group as possible. The most critical descriptive variable is the major occupational classification into which the field test positions were sorted; Tables 15 and 18 indicate that the standardization sample was predominantly technical, service, supervisory, clerical, and managerial in nature. The fact that we obtained high representation of these jobs was deemed desirable, given that these occupational categories have had relatively low representation in the standardization databases collected in many previous general-purpose instruments. A balance between FLSA- exempt versus nonexempt jobs was also obtained (Table 8), which facilitated the process of deriving prediction equations to estimate the FLSA standing of jobs.

With respect to the design goal that the CMQ be completable by job incumbents in a reasonable amount of time, the results in Tables 16-17 indicate that this objective was achieved. Regarding time to complete, approximately two-thirds of the standardization sample completed the CMQ in 2 hours or less, and over 90% of the raters finished in under 3 hours (mode = 1-2 hours). Regarding perceived difficulty of the CMQ rating process, over 87% felt the CMQ rating process was "average" or "easy" in difficulty (mode = "moderately easy").

Scoring the CMQ to Produce Work Dimensions

At the end of the field test, factor analyses were conducted to produce the pools of items used to compute the CMQ's work dimension scores. An 80-scale scoring system was developed; Table 19 presents the coefficient alpha internal consistency reliability estimates for each scale (only alphas for scales computed using 6 or more item ratings are reported).

Table 19. Coefficient alphas for CMQ scales.

#	Dimension Label	rxx
1	Internal Contacts: supervise upper-lvl empls	.82
2	Internal Contacts: supervise mid-lvl empls	.85
3	Internal Contacts: supervise prof/tech empls	.86
4	Internal Contacts: supervise operative empls	.83
5	Internal Contacts: supervise laborers	.83
6	Internal Contacts: marketing/sales empls	.81
7	Internal Contacts: personal service empls	.83
8	Internal Contacts: unions/spec interests	.72
9	Internal Contacts: exchange info/consult	.92
10	Internal Contacts: selling/persuading	.86
11	Internal Contacts: training	.85
12	Internal Contacts: entertaining	.83
13	Internal Contacts: treatment/therapy	.82
14	Internal Contacts: bargaining/negotiating	.81
15	Internal Contacts: resolving conflicts	.79
16	EC: customers/sales related	.86
17	EC: the public/non-sales	.88
18	EC: exchanging business-rel info	.91
19	EC: exchanging info/spec.ints	.88
20	EC: supervision/business-related	.88
21	EC: supervision/suppliers	.84
23	EC: bargaining/negotiating	.83
24	EC: entertaining	.86
25	EC: training	.85
26	CHAIR MTGS: own/other execs	.95
27	CHAIR MTGS: own/other non-exec	.97
28	CHAIR MTGS: nonsupervisory empls	.95
29	CHAIR MTGS: outside supv/exe/pt	.97
30	CHAIR MTGS: public/customers	.92
31	CHAIR MTGS: regul/govt/press	.95
32	CHAIR MTGS: bargain/persuade	.93
33	ATTEND MTGS: own/other execs	.94
34	ATTEND MTGS: own/other non-exec	.96
35	ATTEND MTGS: own/oth/out nonsupv	.91
36	ATTEND MTGS: outside non/sup/pt/exe	.96
37	ATTEND MTGS: public/customers	.91
38	ATTEND MTGS: regulators/govt/press	.95
39	ATTEND MTGS: bargain/persuade	.94
40	FREQ: Internal Contacts/white collar	.81
41	FREQ: Internal Contacts/blue collar	--
42	FREQ: Customers/marketing/sales	--

43	FREQ: EC/public/press/govt/civic	.81
44	FREQ: EC/business-related	.81
45	FREQ: chairing meetings	.93
46	FREQ: attending meetings	.90
47	FREQ: persuading/bargaining	--
48	MDM: POM	.93
49	MDM: Human Resource Responsibility	.93
50	MDM: finance/purchasing-budgeting	--
51	MDM: finance/investments-cash	--
52	MDM: strategic planning/prods-svcs	.85
53	MDM: strategic planning/entire bus	--
54	FREQ: MDM/POM-HRM	.86
55	FREQ: MDM/planning/prod-svcs	.81
56	FREQ: MDM/planning/whole orgs	--
57	FREQ: MDM/HRM/authority-benefits	--
58	FREQ: MDM/finance/budget-purchase	--
59	THINGS: hand-held tools	.82
60	THINGS: stat hand/manuf machines	.82
61	THINGS: vehicle/machine repair	.82
62	THINGS: directing tool/mach use	.84
63	THINGS: office equip/machines	.79
64	THINGS: tech/scientific equip	--
65	THINGS: hvy highway vehicles	--
66	THINGS: lt. highway vehicles	--
67	THINGS: mobile tools/equip	.81
68	THINGS: offroad/utility vehs	.74
69	THINGS: water vehicles	--
70	THINGS: air vehicles	--
71	THINGS: firearms	--
72	FREQ: tool/manuf mach/vehicles	.81
73	FREQ: office machines	--
74	FREQ: tech/scientific machines	--
75	Gross physical activity	.93
76	Precise physical activity	.83
77	Sensory input from work envir	.88
78	Numerical/verbal activ	.75
79	Unpleas/hazardous work envir	.88
80	Demanding personal situations	.81

Procedurally, the CMQ scales were developed by conducting "division" factor analyses of the CMQ item pools. Given the very large number of item ratings that could be used (over 1,200), it was not practical to factor analyze the total pool of CMQ items in a single analysis. Accordingly, each of the main sections of the CMQ (i.e., External

Contacts, Internal Contacts, Managerial Decisions, Meetings Attended, Meetings Chaired, and a final section composed of Things, Machines, and Working Conditions) were analyzed separately, yielding item pools that ranged from 41-180 items in each division.

Frequency item ratings formed a final pool; these were kept separate from the dichotomous item pools in order to separate the kinds of work activity from the frequency with which each is performed. The final 80-factor solution contains 15 internal contacts factors, 10 external contacts factors, 7 meeting chaired factors, 7 meetings attended factors, 6 managerial decision factors, 13 things factors, 6 work context factors, and 16 frequency factors that span the above areas. Refer to the Guide to Using the CMQ for Human Resource Applications (1993) for further details on the CMQ work dimensions.

Factor analyses were conducted using common factor analysis, squared multiple correlation estimates of communalities, eigenvalue plots to determine the number of factors to rotate, and oblique Harris-Kaiser orthoblique factor rotation to identify item-scale linkages. Typically, several eigenvalue discontinuities were observed for each item pool; assessments of the interpretability of each rotated solution were used to decide which solution to use to form the final item pools for each CMQ work dimension scale. When forming these item pools, all items having strong loadings on each factor (using the root-mean-square factor loading as a guide for determining a strong loading) were retained.

In a break from the practices used in many earlier worker- oriented instruments, it was decided not to use a factor analytic scoring system for the CMQ. The reasons for this decision centered on the problems associated with interpreting z scores and in identifying DNA dimensions. That is, "real" factor score estimates are typically scaled in standard (z) units; in a Normal distribution, most of these scores would lie from approximately -3 to +3 z units. The problem with z scores is that they make it very difficult to determine when a work dimension is DNA for a given job: i.e., even if none of the important items defining a factor are rated as being applicable for a given job, a numeric score will still be computed for that job. Thus, it is next to impossible for the typical user of such an instrument to determine whether a given job's dimension score is DNA, or simply reflective of a low (but applicable) level of the dimension.

As the results in Table 18 indicate, the 80 CMQ scales demonstrate acceptable levels of internal-consistency reliability (the median coefficient alpha computed across all scales having 6 or more items was .86). Of course, for a job analysis instrument, one is typically much less concerned with item homogeneity than in other testing situations. That is, the main issues in job analysis are (a) the accuracy of the item ratings, (b) the degree to which the items in the instrument demonstrate content validity in describing the job in question, and (c) the degree to which the scale scores are capable of demonstrating criterion related validity. Nevertheless, these reliabilities support the potential usefulness of the 80-scale system.

The scoring system developed for the CMQ is simple: each scale score is an unweighted linear composite formed from the ratings of the items in each scale's pool. To make the scale scores lie on a similar metric across the different work dimension scales, each raw composite score is converted to a 0-100 metric by dividing by the maximum total number of points possible for each scale. Thus, DNA dimensions are immediately obvious (i.e., 0), unlike the difficulties faced when using a factor analytic scoring system. Values in the 1-100 range indicate the relative percent of the total possible rating points that could have been rated for the position in question.

Of course, the total number of points possible varies across dimensions, so magnitude-based comparisons across different dimensions using these scale scores must only be made after proper consideration of this fact. That is, because these are not true ratio scales, the fact that Job A receives a 30 on dimension 1 and 60 on dimension 2 does not prove that the job requires twice the amount of dimension 2 than dimension 1! Numerical comparisons across jobs should be performed by comparison of the ratings of the same dimensions, each considered separately. For example, it would be proper to conclude that if Job B receives a 50 on dimension 1, then Job B involves a higher level of that dimension than Job A.

To address normative measurement issues, the percentile ranks for the work dimensions are also provided; these were computed using the sample of CMQs collected in the field test. Percentiles are useful when interpreting CMQ profiles, and in making comparisons across positions. Because of the non-Normal distribution of ranked variables, however, such scores have limited use in other settings. See The Guide to Using the CMQ for Human Resource

Applications for further discussion of the CMQ scoring system and the ways in which dimension scores and their percentile ranks should be interpreted.

The CMQ percentiles are unusual in that the rank for each dimension is computed with respect to only those positions in the standardization database that perform the dimension. Thus, the ranks denote the relative ranking of the position relative to the population of jobs that perform the dimension, not the total universe of jobs contained in the standardization sample. It was felt that this method of computing percentiles would have conceptual advantages with respect to giving a normative interpretation of each dimension's score. That is, for work dimensions that are performed by a relatively small percentage of jobs (e.g., high-level managerial decisions), even a relatively low scale score on such a dimension might translate into a very high percentile rank.

One of the major issues confronting a standardized job analysis instrument (other than content validity, of course) is the degree to which scores on its work dimension scales are capable of predicting external variables of interest (i.e., criterion-related validity). Two of the most common external criterion measures are job compensation rates, and whether the job is exempt from the provisions of the Fair Labor Standards Act (FLSA). The CMQ offers scoring services to predict both of these external criterion measures.

With respect to predicting pay, many existing worker- oriented job analysis instruments offer scoring services that predict market pay values using the work dimension scores produced for each job as the predictors. That is, the worker-oriented job dimension scores serve as the compensable factors in such models. Unfortunately, most of these services suffer from the limitation that the "point values" (i.e., predicted compensation rates) they produce are scaled in arbitrary units that do not match the rates of compensation seen in the present-day labor market.

For example, if the pay prediction equation used by a given instrument was derived in 1978 using monthly salary (in 1978 dollars) as the criterion measure, the pay predictions made today using that prediction equation would still predict each job's compensation rate in 1978 dollars. There are at least two drawbacks to this practice: (a) due to the effects of inflation, 1978 dollars form a considerably different metric than 1993 dollars (thus, at best, only the relative ranking of jobs is denoted by the predicted point values); perhaps of greater importance, (b) the kinds of pay policies operative in the general external labor market may have changed in important ways over the intervening 15 years between the derivation of the equation and the present use of the equation to predict wage rates.

Although the former problem of scaling is something of an inconvenience, it might be argued that one can simply transform the dollar metric of the arbitrarily scaled point predictions to match current market conditions by using a linear transformation to take into account the effects of inflation or local market conditions. This is indeed what must be done when arbitrarily scaled points are used to develop a new pay grade system. Assuming that one can identify the appropriate mean and standard deviation that should be present in the organization's new pay structure, this practice may produce acceptable results.

The latter problem of time-based changes in market pay policies is potentially a far more important one. That is, the weights used to define the relative importance of each compensable factor as a predictor of market compensation rates may change considerably over time in response to changes in the external market. Thus, over time we might expect different kinds of work activities to become differentially valued by the market.

For example, relative to the market conditions in the 1970's, today there are significantly fewer unionized workers, and collective bargaining agreements are far more concerned with achieving job security and changing work rules than with achieving significant pay increases for unionized workers. As a consequence, a pay prediction equation developed in the 1970's might contain significantly different weights for job activities that involve physical labor, skilled machine operation, and tool/equipment use than would an equation developed using today's market conditions to define the pay criterion.

Accordingly, the pay prediction system used for the CMQ was designed to be able to predict current market wage conditions, and to adapt to market conditions as they change both over time and over geographic areas. To do this, (a) we used the data from Fall 1992 market wage surveys as the criterion when deriving the initial CMQ market pay prediction equation; (b) we frequently revise this pay prediction formula using updated market wage survey data so that current pay predictions will parallel current market conditions; and (c) we allow CMQ users to request that geographic adjustments (based on the location of the organization in question) be applied to the pay predictions.

The geographic adjustments are also updated frequently to parallel current market conditions.

There are two basic ways in which a pay-prediction equation can be derived: (a) a "local policy capturing" analysis, which uses existing organizational pay data and attempts to derive a system that is unique to each organization to "capture" and perpetuate the pay policies operative in that specific organization; and (b) an external, "market capturing" design that uses market- survey wages for each job and attempts to obtain the best prediction of what a given job would be paid in the external labor market, independent of the current pay practices of the organization.

The CMQ scoring service offers both options; however, unless an organization includes a very large number of jobs in the analysis, it is likely that the most stable and useful results will be obtained by using the national "market capturing" equation in preference to deriving a customized prediction equation.

When deriving the CMQ national "market capturing" pay prediction equation, we used data obtained from current market wage surveys covering hundreds of benchmark jobs as the criterion measure; these market pay averages were matched -- by job title -- to the positions contained in the CMQ field test database. Before deriving the regression equation, an attempt was made to eliminate positions having a high rate of questionable responses; this was accomplished by examining the distributions of the various error counts produced by the CMQ scoring system (see The Guide to Using the CMQ for Human Resource Applications for further details on these indices), as well as the univariate distributions of such items as budget size and the numbers of employees supervised. CMQ profiles that contained suspicious or extreme values were excluded.

Following this exclusionary process, median profiles were formed for each job title with respect to the 80 CMQ job dimension scores and several other single-item fields (e.g., numbers of employees supervised, budget size, work clothing, protective gear, amount of travel, seasonal work). A total of 537 job-median profiles that could be matched to the wage survey database were produced by this process; regarding exempt status, 739 job-median profiles were produced that had FLSA exempt-status data.

Given that a number of the CMQ work dimensions did not exhibit strong bivariate rs with the pay criterion, and in an attempt to keep the number of retained predictors to a minimum, we conducted stepwise regression analyses to identify the final pool of predictors to be used in the national pay prediction equation.

The second kind of external criterion of concern to the CMQ was FLSA-exempt status. That is, a method for being able to estimate the likelihood that a given job or position would be classified as "exempt" has practical utility with respect to both determining a position's initial FLSA classification, as well as in evaluating the validity of an existing classification system.

Although in practice it is perhaps not common to find positions that perform high-level managerial functions being mistakenly classified as "non-exempt," it is not at all uncommon to find lower-level positions being questionably classified as having "exempt" status (e.g., as a means of bestowing additional prestige on a job).

As with the market pay criterion, a number of the CMQ work dimension scores were found to have low bivariate validities with respect to the exempt-status criterion measure; thus, stepwise regressions were conducted to derive the final pool of items to be used for the exempt- status prediction equation.

In both the pay and exempt-status prediction regression models we did not obtain an extremely high ratio of observations to predictors; however, the fact that each predictor profile was computed as the median based on the multiple positions holding a common job title was expected to cause the resultant prediction equations to be much more stable than they would have been if individual positions had been used as the unit of analysis. Thus, although a much higher subjects-to-predictors ratio would have resulted if position -- as opposed to job-median -- profiles had been used, it was decided that the increased stability of the individual predictor profiles that would result from using job medians would be preferable to simply increasing the apparent N used to derived the models.

A stepwise procedure that attempted to find the best level of criterion prediction at each possible number of predictors was used; in this technique, predictors were added and selectively discarded until the highest R for each number of predictors was obtained. To determine the best number of predictors to retain, the multiple Rs produced by each model were plotted against the number of predictors used in the model; inspection of these plots led to the conclusion that a 30-predictor model for market pay (R = .86), and a 19-predictor model for exempt status (R = .74), offered the best balance between predictability and using the minimum number of predictors.

For the pay regression model, the most predictive dimensions included those dealing with managerial decision making, higher level supervision, resource responsibility, business-related travel, negotiation, conflict resolution, and similar work activities. For the exempt- status prediction model, similar dimensions were found to be predictive, with the addition of a few dimensions (e.g., tool and machine use; blue-collar contacts; sales activity) that were not especially predictive of pay. In effect, dimensions of this sort function as inverse predictors of exempt status, and including them facilitates the prediction of nonexempt status.

The R = .86 obtained for the pay prediction equation is comparable to results reported by other worker-oriented instruments (e.g., McCormick et al., 1972; Harvey et al., 1986; Robinson, Wahlstrom, & Mecham, 1974). Although we were unaware of published data reporting Rs obtained when other instruments were used to predict exempt status, the fact that the R for the CMQ exempt- status model was lower than the R for the pay equation led us to speculate that the dichotomous nature of the exempt- status criterion might cause its model R to be an underestimate the degree to which it predicts exempt status.

To assess this possibility, a discriminant analysis was conducted to predict exempt status from the 19 predictors retained in the regression mode; hit rates predicting exempt and nonexempt status were then examined to determine the "bottom line" predictive accuracy obtained using that model. Table 21 presents the hit/miss table for the 19-predictor model obtained using the discriminant functions; the "hits" are located along the diagonal, and "misses" (i.e., cases in which the actual exempt status differs from the predicted status) are off-diagonal. The "rows" of this table correspond to the category into which the job was currently classified; the "columns" of the table represent the classification into which the job was placed by the prediction equation.

Table 21. Hit and miss rates for the nationally derived exempt status prediction model (hits are italicized)

	GROUP THEY WERE CLASSIFIED INTO:	
EXISTING Classification	0	1
Nonexempt 0	*285* *87.42%*	41 12.58%
Exempt 1	68 16.46%	*345* *83.54%*

As the results presented in Table 21 indicate, high hit rates were obtained for both the exempt (84%) and nonexempt (87%) categories using the 19-predictor model. Also as was expected (i.e., based on the above discussion regarding the relatively more frequent case of false "exempt" assignments being awarded), examination of the rates of differentially classified observations in Table 21 indicates that the miss rates were higher for jobs currently classified as exempt than those currently classified as nonexempt. That is, we expected that "truly incorrect" FLSA classifications would more likely be of the variety that falsely classified a job into the "exempt" category than into the "nonexempt" category.

To assess the degree to which the 19-predictor exempt status prediction model produced results that were comparable to those that would be obtained if all predictors were used, a discriminant analysis was also conducted using the full profile of CMQ job dimensions (and associated single-item scales discussed earlier) as predictors of exempt status. Using the model that incorporated all possible CMQ dimensions as the base for the comparisons, the hit rates produced by that model were within 1-2% of the results reported in Table 21 for the model using only 19 predictors. Accordingly, the relatively trivial increase in predictive power made possible by using far more predictor dimensions in this second model was not deemed to offset the expected decrease in stability caused by using more predictors.

Finally, for illustrative purposes, this paper reports the results of "local policy capturing" analyses performed using the CMQ scores as predictors of pay and exempt status. That is, instead of using the above-described national market-policy pay prediction equation to generate pay predictions, organizations may opt to have the CMQ scoring service derive a prediction model that uses the organization's own compensation rates and exempt status classifications as the criterion measures (i.e., a "local policy capturing" study).

Obviously, the results obtained in such an analysis will be specific to the individual organization, and quite different results may be obtained in different organizations. Keeping this fact in mind, however, it is nevertheless interesting to examine the results of such "local policy capturing" studies using several organizations drawn from the CMQ field-test database; the actual reported pay (converted to yearly dollars) and exempt status for each position served as the criteria. In effect, these analyses illustrate what would have happened had these CMQ clients requested a "local policy capturing" analysis be conducted.

For each organization, all CMQ scales considered for use in the above national equations (less any predictors that possessed no variance in each organization) served as predictors. The organizations selected were as follows: (a) 226 positions (ranging from night watchman to CEO; 44.8% currently classified as exempt) from the home office of a life insurance company; (b) 272 positions from a hospital (entry level to department head level; 11.4% exempt); (c) 203 positions from a computer hardware manufacturing company (all employees; 70.0% exempt); (d) 220 positions from a city government (low- to mid- level; 39.5% exempt); and (e) 142 positions from a paper products company (entry- to upper-level; 14.8% exempt).

Table 22 presents the multiple Rs obtained from predicting existing position pay and exempt values in these field-test organizations. Hit rates predicting exempt-status were produced by dichotomizing the predicted criterion value from the regression equation; positions predicted to have a 50% or greater probability of being exempt are classified as exempt. Given the relatively large number of predictors used (relative to the number of positions in each organization), both raw and shrunken Rs are reported.

Table 22. Regression Results Predicting Pay and Exempt Status

Model	N. Profiles	N. Preds.	R	Rs	HitsN	HitsE
Insurance Company, Pay	226	87	.968	.947	--	--
Insurance Company, Exempt	226	87	.888	.810	98.4	92.1
Hospital, Pay	272	89	.846	.760	--	--
Hospital, Exempt	272	89	.843	.754	99.6	80.6
Computer Company, Pay	203	88	.930	.872	--	--
City Government, Pay	220	91	.823	.669	--	--
City Government Exempt	220	91	.770	.551	97.7	73.6
Manufacturing Co., Pay	142	89	.992	.977	--	--
Manufacturing Co., Exempt	142	89	.979	.941	100.0	100.0

Note. Rs = shrunken R; HitsN = hits for nonexempt status; HitsE = percentage of hits for currently exempt employees; N. Preds. = number of predictor scales used in model; N. Profiles = number of position profiles used.

As the results presented in Table 22 indicate, very high rates of prediction of existing position pay and exempt status are possible using a local policy-capturing design: three of the organizations reported multiple Rs in the mid-to high .90's when predicting pay, and all organizations reported hit rates predicting nonexempt status in the high 90% range. As with the national sample, hit rates were somewhat lower predicting exempt status, and considerable variability existed across organizations (e.g., the city government exempt status classification was much less predictable).

Of course, this high rate of prediction comes at a cost: namely, possible instability of the regression equation when applied to new samples, as well as the potential for capturing -- and perpetuating -- both desirable and undesirable sources of variance in the criterion (especially pay). Regarding the use of locally provided pay rates as criterion measures in a "local policy capturing" study, the possibility that systematic biases in pay rates (e.g., those that work to the relative disadvantage of female- dominated jobs) might exist must be carefully considered. That is, if the criterion contains such biases, even if it is highly predictable the resulting system might simply function to perpetuate the existing biases, which would clearly be an undesirable outcome.

For these reasons, we encourage all organizational users of the CMQ that are sensitive to external equity concerns to use the nationally derived, frequently updated pay prediction equation in preference to conducting a locally derived policy-capturing analysis. By using the national equation, which was developed using the results of nationwide wage surveys of hundreds of different jobs, organizations can be considerably more confident that (a) the pay criterion data are as free from systematic biases as is possible when market data are used; and (b) that all jobs (regardless of the percentage of male versus female position incumbents) are being treated consistently with respect to the application of pay policies.

I have previously hypothesized (e.g., Harvey, 1985) that although it is possible -- indeed, even likely -- that some quarters of the market have tended to pay female- dominated jobs at lower rates than male-dominated jobs that require comparable skill, effort, responsibility, or working conditions, perhaps the most important source of sex-based pay inequity lies in the process by which each job is rated on the various compensable factors during the job analysis process. That is, sex-based stereotypes may significantly color the ratings given by job analysts or a compensation committee regarding each job's standing on each compensable factor (see Harvey, 1991a, pp. 105-106, 145-146).

Although gender-based pay biases may occur due to many factors, the one area in which something can actually be done to reduce bias involves the initial process of rating the compensable factors. This potential source of bias is especially problematic for job analysis instruments that attempt to holistically rate behaviorally nebulous characteristics of jobs (e.g., Overall Responsibility, Importance of Decisions Made) and then use these ratings to define the compensable factors. That is, (a) research has demonstrated that holistic ratings of vague job traits have highly questionable psychometric properties and accuracy (e.g., Butler & Harvey, 1986; Harvey, Wilson, & Blunt, 1993); and (b) ratings of behaviorally vague traits are effectively impossible to verify independently (i.e., because they do not describe aspects of the job that are observable and therefore subject to independent evaluation).

Summary & Future Directions

The CMQ was developed to address a number of potentially serious shortcomings that limit many existing worker-oriented job analysis questionnaires. Specifically, it was designed to (a) be easier to comprehend, and therefore be usable by average job incumbents; (b) produce ratings that are more behaviorally specific, and therefore verifiable; and (c) span the content domain of both exempt and nonexempt jobs, and therefore allow meaningful comparisons between all kinds of jobs.

Based on the results of the CMQ field test described above, it seems reasonable to conclude that these objectives were achieved. Regarding comprehensibility and usability, the eighth-grade reading level target that was set produced an instrument that was completed by job incumbents in most of the field-test organizations. These raters overwhelmingly judged the CMQ to not be difficult to complete, and over 90% completed it in 2-3 hours or less.

Regarding the goal of providing a more behaviorally detailed description of work, the CMQ produces over ten times the number of item ratings (2,077 versus 194), and approximately twice the number of work dimension scores (80 versus 32 division plus 13 overall dimensions) than the PAQ, the instrument that has in the past been viewed as the standard general-purpose instrument. Overall, the CMQ's items are much more behaviorally specific -- and therefore, verifiable -- than many of the extremely abstract items contained in many worker-oriented instruments. In the field test sites that conducted systematic reviews of the incumbent-completed CMQ booklets, these reviews were greatly facilitated by virtue of this higher level of behavioral specificity in the CMQ items.

With respect to the goal of being able to analyze both exempt and nonexempt jobs, a review of the dimension targets that were set for the CMQ, and the work dimension scores produced by the CMQ scoring system, reveals that clear progress was made toward achieving the goal of increasing the number -- and specificity -- of the job dimensions that describe the work activities performed on higher-level jobs. Although any assessment of the content validity of the CMQ's item pool, or the construct validity of its work dimensions, is by definition subjective, it seems reasonable to conclude that the CMQ is at least as descriptive of the work activities performed in nonmanagerial jobs -- and much more descriptive of those performed in professional and managerial jobs -- than earlier general-purpose instruments like the JEI and PAQ. Consequently, the CMQ should be much more useful for purposes that depend on being able to make meaningful comparisons between diverse collections of jobs or occupations.

Of course, additional work remains to be performed. Although the initial pay-prediction equation produced highly satisfactory results, efforts will continue to further increase the levels of pay prediction that are produced. Likewise, research will continue to derive equations to predict the ability/aptitude requirements of jobs using job component validity techniques; once completed, these empirically derived linkages will make it much easier to estimate the ability requirements of jobs based on knowledge of the required worker activities.

A highly promising future direction for the CMQ lies in the area of integrated personnel systems (IPS; e.g., Wilson, 1987). The key notion of the IPS is that a job analysis database lies at the heart of any effective personnel system, and the various personnel functions that make use of job analysis information (e.g., selection, performance appraisal, compensation) should be tightly linked to this central database (ideally, via automated computer technology). Although it is obvious that a single job analysis instrument cannot possibly serve every possible personnel function, the CMQ has a number of properties that make it an appealing choice to serve as the foundation for the centralized job analysis database in an IPS (e.g., Harvey, 1993).

That is, the CMQ's ability to describe both exempt and nonexempt jobs using a common profile of GWBs -- combined with its higher level of behavioral specificity than is typically seen in worker-oriented questionnaires -- render it useful for a wide variety of functions (e.g., performance appraisal, job descriptions, compensation, job classification, career path planning, skills transferability, test validation, synthetic validity). Additional research (e.g., Harvey, 1993) is now underway to develop computer-aided methods for linking the CMQ profiles included in an IPS job analysis database with a much larger database of task statements. Under such a system, once a job's CMQ profile is known, an interactive computer program can guide a job incumbent (or other SME) through a series of questions designed to identify a listing of tasks that are performed on the job, without the need to conduct a separate task analysis.

In effect, the moderate level of behavioral specificity contained in the CMQ items allows it to function as a "linking pin" between the domain of abstract work dimensions and required worker ability traits (when moving in the more behaviorally abstract direction), as well as the domain of job task statements (when moving in the more behaviorally specific direction). Thus, by knowing a job's CMQ item-level profile, more abstract work dimension scores can be produced, and these work dimension scores can be linked (via synthetic validity studies) with the domain of required worker ability traits.
Moving in the other direction, once the CMQ items that are performed on a job are known, interactive computer programs can be used to guide the development of listings of the more detailed task statements that characterize the job (either by linking to existing databases of task statements or by prompting the SME to generate new statements using an expert-system approach).

Thus, the CMQ may finally offer a solution to the age-old problem of the high cost associated with collecting task-level job analysis information. Thus, only the time and cost of administering the CMQs -- plus the SMEs' time that is spent using the interactive computer program to select and/or generate the task statements performed on each job -- would be required to develop a comprehensive task- and worker-oriented job analysis database. Using this approach, it would therefore be cost-effective for even the smallest organizations to develop state-of-the-art integrated personnel systems.

References

Ash, R. A., & Edgell, S. L. (1975). A note on the readability of the Position Analysis Questionnaire (PAQ). *Journal of Applied Psychology, 60*, 765-766.

Ash, R. A., & Levine, E. L. (1980). A framework for evaluating job analysis methods. *Personnel, 57*, 53-59.

Cornelius, E. T., & Hakel, M. D. (1978). *A study to develop an improved enlisted performance evaluation system for the U.S. Coast Guard*. Washington, DC: Department of Transportation.

Cunningham, J. W. (1964). *Worker-oriented job variables: Their factor structure and use in determining job requirements*. Unpublished doctoral dissertation, Purdue University.

Cunningham, J. W., & Ballentine, R. D. (1982). *The general work inventory*. Raleigh, NC: Authors.

Cunningham, J. W., Boese, R. R., Neeb, R. W., & Pass, J. J. (1983). Systematically derived work dimensions: Factor analysis of the OAI. *Journal of Applied Psychology, 68*, 232-252.

Fine, S. A. (1989). *Functional job analysis scales: A desk aid*. Milwaukee, WI: Author.

Harvey, R. J. (1985, November). *Comparable worth compensation systems: Implications for the bottom line*. In R. J. Harvey (Chair), *Comparable worth: Employment issue of the 80's?* Symposium presented at the Texas Psychological Association Convention, Houston.

Harvey, R. J. (1987, April). *Alternative factor structures for the position analysis questionnaire (PAQ).* In M. D. Hakel (Chair), *The dimensionality of work: Future directions, applications, and instrumentation*. Symposium presented at the annual conference of the Society for Industrial and Organizational Psychology, Atlanta.

Harvey, R. J. (1990). *The common-metric questionnaire for the analysis and evaluation of jobs (field test version 1.12)*. San Antonio, TX: The Psychological Corporation.

Harvey, R. J. (1991a). *Job analysis*. In M. D. Dunnette and L. Hough (Eds.), *Handbook of industrial and organization psychology* (second edition). Palo Alto, CA: Consulting Psychologists Press.

Harvey, R. J. (1991b). *The common-metric questionnaire (CMQ): A job analysis system* (first edition). San Antonio, TX: The Psychological Corporation.

Harvey, R. J. (1993, May). *The common-metric questionnaire: Applications and current research*. In S. Fine (Chair), *Multipurpose job analysis: New approaches supporting integrated human resource management*. Symposium presented at the annual conference of the Society for Industrial and Organizational Psychology, San Francisco.

Harvey, R. J., Friedman, L., Hakel, M. D., & Cornelius, E. T. (1988). Dimensionality of the Job Element Inventory, a simplified worker-oriented job analysis questionnaire. *Journal of Applied Psychology, 73*, 639-646.

Harvey, R. J., & Hayes, T. L. (1986). Monte carlo baselines for interrater reliability correlations using the Position Analysis Questionnaire. *Personnel Psychology, 39*, 345-357.

Harvey, R. J., Wilson, M. A., & Blunt, J. H. (1993). A comparison of rational/holistic versus empirical/decomposed methods of identifying and rating general work behaviors. Manuscript under review.

Jeanneret, P. R. (1969). *A study of the job dimensions of "worker-oriented" job variables and of their attribute profiles*. Unpublished doctoral dissertation, Purdue University.

Lozada-Larsen, S. R. (1986). *Uncovering the Dimensionality of Managerial, Supervisory, Executive, and Professional Work*. Unpublished Master's thesis, Rice University.

Lozada-Larsen, S. R. (1988). *Going beyond criticism: Management work theory and research*. Unpublished doctoral dissertation, Rice University.

McCormick, E. J. (1959). The development of processes for indirect or synthetic validity: III. Application of job analysis to indirect validity: A symposium. *Personnel Psychology, 12*, 402-413.

McCormick, E. J. (1976). *Job and task analysis*. In M.D. Dunnette (Ed.), *Handbook of Industrial and Organizational Psychology*, 651-696.

McCormick, E. J., Cunningham, J. W., & Gordon, G. G. (1967). Job dimensions based on factorial analyses of worker-oriented job variables. *Personnel Psychology, 20*, 417-430.

McCormick, E. J., Jeanneret, P. R., & Mecham, R. C. (1972). A study of job characteristics and job dimensions as based on the Position Analysis Questionnaire (PAQ). *Journal of Applied Psychology, 56*, 347- 367.

Mitchell, J. L., & McCormick, E. J. (1979). *Development of the PMPQ: A structural job analysis questionnaire for the study of professional and managerial positions*. (Report No. 1). Lafayette, Ind.: Occupational Research Center, Department of Psychological Studies, Purdue University.

Page, R. C., & Gomez, L. R. (1979). *The development and application of job evaluation systems using the Position Description Questionnaire*. (Personnel

Research Report No. 162-79). Control Data Corporation.

Palmer, G., & McCormick, E. (1961). A factor analysis of job activities. *Journal of Applied Psychology, 45*, 289-294.

Robinson, D. D., Wahlstrom, O. W., & Mecham, R. C. (1974). Comparison of job evaluation methods: A "policy capturing" approach using the Position Analysis Questionnaire. *Journal of Applied Psychology, 59*, 633-637.

The Psychological Corporation (1993). *The guide to using the CMQ for human resource applications*. San Antonio, TX: Author.

Wilson, M. A. (1987, April). *Work dimensionality and integrated personnel systems*. In M. D. Hakel (Chair), *The dimensionality of work: Future directions, applications, and instrumentation*. Symposium presented at the annual conference of the Society for Industrial and Organizational Psychology, Atlanta.