

Why do IQ scores predict job performance? An alternative, sociological explanation

Eliza Byington * & Will Felps

ABSTRACT

During the past century, IQ testing has become a pervasive tool for allocating scarce resources in the United States and beyond. IQ-reflective tests are used in primary and secondary schools to sort students into groups, and by universities and employers to select between applicants. Drawing on several sociological literatures (i.e. the diffusion of institutions, social stratification, and self-fulfilling dynamics), we argue that the current applications of IQ-reflective tests allow individuals with high IQ scores to receive greater access to developmental resources, enabling them to acquire additional capabilities over time, and ultimately perform their jobs better. This institutional explanation for the IQ–job performance link is an alternative to the prevailing assertion that IQ facilitates job performance through on-the-job learning. Further, this explanation suggests that the frequently repeated prescription “select on intelligence” may be institutionally contingent.

Research in Organizational Behavior

Volume 30, 2010, Pages 175-202

Copyright © 2010 Elsevier Ltd All rights reserved.

doi:10.1016/j.riob.2010.08.003

* Corresponding author: E. Byington, ebyington@rsm.nl

Over the past century, numerous studies have documented the link between cognitive assessment scores and employee performance (see Schmidt & Hunter, 1992, 1998, 2004 for overviews). The strength of these findings have made intelligence “the most important trait or construct in all of psychology, and the most ‘successful’ trait in applied psychology” (Schmidt & Hunter, 1986). Few areas of scholarship produce such unequivocal findings and prescriptions for managers. In a chapter advocating the principle that employers should “select on intelligence,” Schmidt and Hunter state: “Intelligence is the major determinant of job performance, and therefore hiring people based on intelligence leads to marked improvements in job performance – improvements that have high economic value to the firm” (2000: 3). The assertion that intelligence leads to job performance is echoed in the titles of recent articles in which management researchers describe their findings: “Intelligence is the best predictor of job performance” (Ree & Earles, 1992), “The role of general cognitive ability and job performance: Why there cannot be a debate” (Schmidt, 2002), and “Predicting job performance: Not much more than g” (Ree, Earles, & Teachout, 1994). Building on this empirical foundation, researchers have gone so far as to provide methods by which managers and HR professionals can calculate the positive economic value of employing IQ-based selection in their organizations (e.g. Rauschenberger & Schmidt, 1987; Schmidt, Hunter, McKenzie, & Muldrow, 1979; Hunter & Schmidt, 1996; Schmidt & Hunter 1998). The industrial psychology literature has also reached a consensus on the explanation for the strong IQ – job performance relationship; namely, that individuals who are more intelligent (as measured by IQ scores) can learn job relevant knowledge faster and better, resulting in improved job performance (e.g. Ree, Carretta, & Teachout, 1995; Hunter, 1986).

In this article, we challenge this explanation by introducing an alternative, sociological model that can provide new insights into why cognitive ability scores predict job performance so

well. Further, we explore how this new understanding can inform managerial practices, as well as highlight important, institutional boundary conditions for generalizing about IQ - job performance findings. To begin, we first offer an overview of the conventional wisdom within industrial psychology about how, why, and when intelligence relates to job performance. In this review, we consider how intelligence has been operationalized, and note the many incarnations of IQ testing used by organizations¹. In the next section, we review the historical process by which IQ tests came to be a widely institutionalized basis for the allocation of resources by organizations. Building on the historical diffusion of IQ testing, we then present an alternative, theoretical model for how, why, and when IQ scores predict job performance. Specifically, we argue that many organizations rely on IQ indicative credentials to distribute developmental resources² that may, in and of themselves, facilitate eventual job performance. We further argue that such institutionalization can result in a self-fulfilling dynamic, whereby a social science technology becomes embedded in institutional designs in ways that increase the validity of the original idea behind the technology (e.g. Ferraro, Pfeffer, & Sutton, 2005). In closing, we discuss how the inclusion of these factors in the dominant industrial psychology model can inform managerial practices, public policy, and research.

¹ In this article, *organization* denotes a social arrangement with common resources that collectively directs its own activities, and has a boundary separating it from its environment. Thus, it includes corporations, universities, schools, and the military.

² Note that we use *resources* as an umbrella term to refer to something that is both scarce *and* instrumentally valued for accomplishing some terminally desired end-state. This broad conception of resources as “scarce + instrumentally valued” is shared within both sociological (Bourdieu, 1986: 248) and managerial (Wernerfelt, 1984) traditions. This article will go on to detail the specific resources at play.

REVIEW OF THE IQ – JOB PERFORMANCE LITERATURE

Defining the Variables

In the intelligence - job performance literature, “intelligence” is measured with a variety of psychometric assessments that are often referred to as “IQ tests”. Common among these are Raven's Progressive Matrices, the Armed Services Vocational Aptitude Battery, Wechsler's Adult Intelligence Scale, and the Wonderlich Intelligence Test. Researchers also use many tests which are not commonly thought of as “IQ tests” as measures of *intelligence* because of their strong correlation with explicit measures of IQ. A number of university admissions tests are highly indicative of an individual's IQ score, including the SAT, ACT, and graduate school entrance exams like the GRE and GMAT. For example, Frey and Detterman (2004) have found that, when corrected for non-linearity, SAT scores correlate at approximately .86 with tests explicitly designed to measure IQ. As a consequence, such tests can also be used to link the psychometric characteristic of IQ to individual outcomes. Indeed, the high correlations among the variety of cognitive assessments, factor analysis, and overlapping predictive validities have been used to suggest that these assessments are all, to varying degrees, measuring the underlying trait “g” – Spearman's *general intelligence factor* (e.g. Ree et al., 1994; Rindermann, 2007a).

Both in common usage and the job performance literature, the terms *IQ* and *intelligence* are often used interchangeably. However, to avoid conceptual confusion, it is important to note that IQ scores are artifacts of a particular sort of technology (IQ tests), which do a better or worse job of assessing the construct *intelligence*. While there is little debate that a common factor underlies scores on the variety of above mentioned tests, there is considerable discussion as to whether this factor should be called *intelligence*. Many suggest that intelligence is a broader or different phenomenon (for diverse perspectives on this debate, compare Flynn, 1999; Gardner, 1999; Jensen, 1998; and Sternberg, 1988).

For the arguments developed in this article, we are agnostic on the question of how well IQ tests measure intelligence. *Regardless* of the extent to which IQ scores reflect real differences in intelligence, the theoretical model we develop suggests that differences in IQ scores (and IQ related scores, and credentials) will drive the allocation of developmental resources and ultimately job performance. Thus, we do not endorse any particular conception of intelligence, but are instead interested in the consequences of IQ scores, IQ-reflective scores³, and IQ indicative credentials for job performance. Finally, we define job performance as the degree to which episodes of employee behavior improve organizational effectiveness (see Campbell, Gasser, & Oswald, 1996 for a similar definition).

The Conventional Wisdom about IQ and Job Performance

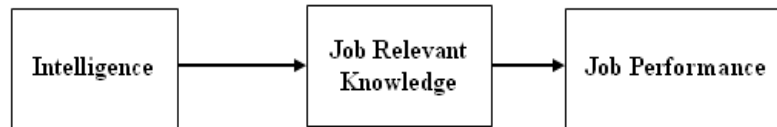
In the industrial psychology literature, IQ scores are most often used to predict job performance.⁴ According to Schmidt & Hunter, “[l]iterally thousands of studies [investigating the IQ – job performance relationship] have been conducted over the last nine decades” (1998: 264). These studies have been summarized in a number of meta-analyses (Grudnik & Kranzler, 2001; Hunter, 1986; Kuncel Hezlett, & Ones, 2004; Levine, Spector, Menon, Narayanan, & Cannon-Bowers, 1996; McHenry, Hough, Toquam, Hanson, & Ashworth, 1990; Schmidt & Hunter, 1998). Recent empirical analysis on the relationship between IQ and job performance has been used to estimate an overall correlation of between .45 and .58 (Schmidt & Hunter, 2000: 4). Job complexity acts as a key moderator of this relationship, with stronger effects for more complex jobs (Gottfredson, 2002).

³ The term “IQ-reflective” tests to refer to those assessments that a) correlate highly with IQ tests, have b) the same standardized, multiple-choice format, and c) the aim of assessing intellectual ability.

⁴ A few studies have also examined links between IQ scores and citizenship behaviors (e.g. Cote & Miners, 2006), workplace deviance (e.g. Dilchert, Ones, Davis, & Rostow, 2007), and job satisfaction (e.g. Ganzach, 1998).

To explain these relationships, industrial psychologists have offered a straightforward theoretical model. It is argued that, “The main reason that general cognitive ability predicts job performance is because it predicts how well an individual learns job relevant knowledge. That is, general cognitive ability predicts the learning of the job” (Hunter, 1986: 359). The traditional industrial psychology model is summarized in Figure 1.

FIGURE 1

The Conventional Industrial Psychology Model*

*Adapted from Hunter, 1986.

Researchers have provided evidence of statistical mediation of the IQ - job performance relationship by “job knowledge” on a number of occasions (see Hunter, 1986 for a meta-analysis). Perhaps the most rigorous test to date has been conducted by Ree, Carretta, and Teachout (1995), who find full mediation of the IQ – performance link by job knowledge among their sample of pilots in training.

Schmidt and Hunter argue that: “If a group of individuals were to begin a job new to them at the same time, then there would be no experience differences and experience could not have an effect on individual differences” (1992: 90). We reject that assumption, and instead argue that much job-relevant knowledge, as well as the skills necessary to acquire specific “job knowledge”, are in fact capabilities developed prior to a particular job. This rejection is not radical. Labor economists studying “human capital” describe a variety of types of knowledge which can be used

to improve work performance (for an introduction, see Becker, 1993). These typologies include job specific knowledge, but also recognize occupation specific (Kambourov & Manovskii, 2009) and industry specific knowledge (Neal, 1995). Indeed, a variety of studies by labor economists indicate that knowledge learned at one job is usually quite portable to subsequent jobs (e.g. Lillard & Tan, 1992). Moreover, much of the knowledge learned in school is general knowledge that can apply to many sorts of work. General knowledge, for example, includes generic work skills like reading comprehension and arithmetic, as well as metacognitive skills such as: a) recognition, definition, and representation of problems, b) the ability to monitor task progress and change strategies when necessary, and c) evaluating solutions (Becker, 1993). Labor economists have shown that there are sizable returns on investments in such general knowledge (Blundell, Dearden, Meghir, & Sianesi, 1999). Beyond labor economics, the more inclusive notion that job relevant knowledge is acquired throughout one's life is consistent with the literatures on learning, ability, and expertise in psychology and education (e.g. Sternberg, 1998; Gladwell, 2008). Research within the industrial psychology tradition has not attempted to disentangle the sources of job relevant knowledge, proceeding instead under the simplifying assumption that it is acquired on-the-job (e.g. Hunter, 1986; Ree et al., 1995).

Moreover, a second critique involves the meta-analytic techniques used to establish the generalizability of the IQ – job performance link. Claims that employers should “select on intelligence,” and that “there are no known exceptions to this principle” (Schmidt & Hunter, 2000: 7) are premised on the results of the validity generalization (VG) model developed by Schmidt and Hunter (1977; see Hunter & Schmidt, 2004 for an update on VG practices). VG seeks to correct for sampling error, range restriction, and unreliability in measures. With respect to IQ, validity generalization has suggested that, at any given level of complexity, the correlation between IQ and performance is quite similar across jobs, once the appropriate corrections have

been made (Hunter, 1986). This VG evidence was used to refute earlier claims that the IQ – job performance relationship might only pertain to specific types of work.

However, the validity generalization program has recently come under scrutiny from scholars who question some of the assumptions built into the model (e.g. Murphy, 2003, 2009; Schmitt, 2007). One of the most serious of these critiques is that the method assumes that the studies on which a VG analysis is based were randomly drawn from the population of interest. This issue has raised “a fundamental question that is rarely addressed in VG studies—i.e., defining what is meant by ‘the population’” (Murphy, 2009: 443). Thus, the results produced by “validity generalization” can only be appropriately generalized to contexts similar to those that have been included in the VG analysis. In the case of IQ, such generalizations are problematic. While an enormous number of studies have been done on IQ and job performance, the vast majority come from North American samples. The first meta-analyses of studies from outside North America were recently conducted by Salgado and colleagues (2003a; 2003b). These two meta-analyses drew on basically the same samples, specifically studies from ten European countries (the majority coming from Northern European countries such as the Netherlands, Germany, and the United Kingdom). They find that “[t]he magnitude of the operational validity [of IQ in predicting job performance] for the European studies is remarkably larger than the magnitude reported in the American meta-analyses” (2003a: 580). These results undermine the claim of universally generalizable effect sizes.⁵ Moreover, even these meta-analytic studies are from highly developed nations.

⁵ Later in the article, we discuss why these effects may be stronger (rather than weaker) in a Northern-European sample.

In preparing this article, we engaged in an exhaustive search of published studies connecting IQ and job performance in developing countries (e.g. China, India, Malaysia, etc.). This involved using the ISI Web of Science to search for abstracts which included IQ (or ability, intelligence, cognitive, GMA, or “g”) and job performance (or task performance) for every developing country (as designated by the International Monetary Fund). In addition, we had a colleague search Chinese language databases for articles on these topics. These extensive searches yielded surprisingly few relevant articles. The first article, by Law and colleagues (Law, Wong, Huang, & Li, 2008), found a statistically insignificant *negative* relationship ($r = -.07$) between IQ and job performance in a sample of research and development professionals in China. A second article by Reeb (1976) compared the relationship between IQ and job performance for soldiers in the Israeli Army. The IQ – job performance relationship was considerably stronger for soldiers from Israel ($r = .43$) than for Israeli soldiers from the Middle East or West Africa ($r = .28$). Reeb suggests that the substantial differences in IQ – job performance correlations between these two groups of soldiers might be due to their different levels of education. The results of this expansive search are suggestive that empirical studies on this subject do not exist in abundance outside of first-world countries.

In sum, our suggestion is that the strength of the IQ – job performance relationship may not be universal given a) the unclear source of “job knowledge”, b) concerns about the randomness of samples used in VG analysis, and c) the differing magnitude of the effect size outside of first-world nations in the few empirical studies that have been conducted. The following section explores how varying degrees of institutionalization of the rule “select on IQ” may limit the generalizability of IQ – job performance claims.

A Sociological Alternative

Our article makes the claim that both within and outside of the workplace, IQ scores have become embedded in institutional designs as a criterion for the allocation of scarce developmental resources. If IQ testing had failed to become embedded in institutional designs, we argue that the IQ score – job performance relationship might be appreciably smaller. This point is important for how much managers, particularly those in non-American contexts with very different institutional designs, should rely on the prescription “select on intelligence”. Further, we underscore the prevalence of IQ testing as a path-dependent, historical accomplishment (Mahoney, 2000). In particular, the expansion of IQ-testing into American organizational practices reflects the contingencies of two world wars, the advocacy of a few key institutional entrepreneurs, and fit with cultural and bureaucratic ideals. It should be noted that while the institutionalization of IQ testing has spread far beyond America, the initial diffusion is generally seen as an American phenomenon (e.g. Carson, 2007). Thus, the following section provides a historical case study in how American organizations adopted IQ-reflective testing during the 20th century.

THE HISTORICAL INSTITUTIONALIZATION OF IQ TESTING

When studying the process of institutionalization, historical research can provide insights into how and why managerial practices diffuse. Ferraro, Pfeffer, and Sutton note the need for incorporating history into theory:

“Rich historical research on the archeology or origin of management practices is needed to understand the longitudinal process of transforming theories into management practices, as well as the factors that have led designers and adopters to choose some practices over others... More work in this vein is

needed on management practices to better specify the relationship between social science and practice” (2005: 21).

Many characteristics of IQ tests make them particularly amenable to organizational application. Indeed, from the beginning, IQ testing was designed for use by bureaucratic institutions (Siegler, 1992). In this section, we review the origins of IQ testing, and the historical contingencies that lead to the diffusion of IQ-reflective testing across a broad range of organizations, including the Army, universities, schools, and firms in 20th century America. In particular, we highlight the factors that hastened the spread of testing into these new contexts and the generalization of this technology to an expanding set of organizational objectives. The widespread diffusion of IQ-reflective tests across organizations is a precondition to the theoretical model that follows this section.

The Beginnings of the Science of Intelligence

Philosophers have long been interested in human intelligence, but the first scientifically rigorous attempts to examine intelligence empirically are credited to Sir Francis Galton (Siegler, 1992). In his investigations, Galton noted that individuals who performed well in one domain tended to also perform well in others. This observation prompted him to hypothesize an underlying characteristic that drove all cognitive performance. To measure this characteristic, Galton devised tests of reaction time as a proxy for mental processing speed (Siegler, 1992).

In the beginning of the 20th century, Alfred Binet extended the study of intelligence to complex mental processes (Siegler, 1992). Binet and his research assistant Henri Simon built on this conceptualization of intelligence to devise a test that would identify children likely to need additional educational support in the new system of compulsory public education being introduced in France at the time. Binet attempted to make performance on his assessment, the first IQ test, as objectively measurable as possible (Siegler, 1992). He fundamentally distrusted

the objectivity of parents and teachers as test administrators, suggesting that parents might not want to admit their children's cognitive deficiencies. Teachers might also be motivated to rate their student's performance highly as it could be considered a reflection on their teaching effectiveness. Alternately, teacher discretion in the interpretation of student performance might be influenced by a student's disciplinary history in class.

As a consequence of these concerns, and the pragmatic needs of the French school system, the test Binet introduced in 1905 was made to be "inexpensive, easy to administer, objectively scorable, and predictive of classroom performance" (Siegler, 1992: 185). These characteristics of the IQ test persist to this day and, as we will review, have played an important role in the widespread diffusion of the practice. Some of the other features of Binet's original scale that have remained include the use of a variety of subtests, questions that involve pattern inference, spatial reasoning, vocabulary and reading comprehension, problems ordered to increase in complexity, and the practice of reporting performance relative to that of other individuals of the same age (Siegler, 1992). Given the prevailing trends in French psychology at the time, however, IQ testing's greatest successes would not be in France, where Binet's test was virtually unknown for many years, but in the United States (Siegler, 1992).

The Diffusion of IQ Testing in America

It was also during the beginning of the 20th century that the newly formed American Psychological Association (APA) began an effort to coordinate the collection of data on individual differences in cognitive performance in the United States (Siegler, 1992). This movement was motivated by the goal of developing psychometric tests that would reliably reveal individual differences in a quantifiable "statement of worth" (Baritz, 1960: 47). Among those active in this endeavor were Henry Goddard, G. Stanley Hall, James McKeen Cattell, E. L. Thorndike, Robert Yerkes, and Lewis Terman. Their shared interest in what could be achieved

through psychometric testing lead to the further refinement of Binet's measures. However, the American public was highly skeptical of the validity of such tests (Kevles, 1968). Kevles highlights the ethos of the day with a quote from Justice John W. Goff of the Superior Court of New York in 1916, who said: "Standardizing the mind is as futile as standardizing electricity," in ruling that IQ tests could not serve as a means of assessing "feeble-mindedness". The public resistance toward intelligence testing makes the widespread diffusion of the practice in the years that followed that much more remarkable. It was in this climate that Stanford psychology professor Lewis Terman introduced his revised version of Binet's test in the United States (Terman, 1916). From these humble beginnings, IQ testing spread across a variety of organizational contexts, from the military, to universities, schools, and firms.

IQ testing in the US Army. Scholarship by Kevles (1968) documents the first large-scale, systematic, organizational adoption of IQ testing in the U.S.. With the outbreak of World War I, APA president Robert Yerkes convened a dozen special committees to consider military applications for the young science of psychology (Kevles, 1968). Initially, "mental testing" was proposed as a criterion for assessing draftees who might have cognitive deficiencies, making them unfit to serve in the military. But Yerkes wanted to extend intelligence testing beyond Binet's emphasis on cognitive deficiencies. He advocated for the use of Terman's revised version of Binet's IQ test as a placement device for all personnel within the military, as well as a selection criterion for fitness to serve.

At the beginning of World War I, the Army's placement decisions for draftees were made on the basis of time-intensive, first-hand observation by commanding officers. For officers, seniority was the primary factor in promotions (Kevles, 1968). These placement criteria considerably restricted the speed with which draftees were processed, and severely limited the number of officers. Given the extreme time pressure to quickly process new recruits, Yerkes was

given approval to implement IQ testing for all “enlisted men, draftees, and officers already appointed” (Kevles, 1968: 571). After occupational experience, IQ scores became the basis for placing draftees in particular companies, training regiments, and jobs (Carson, 2007: 210). Assignments were made such that companies received a set proportion of men with high, average, and low IQ test scores (Kevles, 1968). This process marked a significant transformation within the Army to standardized, psychometric staffing (Bingham, 1942).

The primary reason for the Army’s adoption of IQ tests (then called the Alpha and Beta tests) for selection is clear. Compared to the pre-war methods (i.e. qualitative assessment and seniority), IQ-based selection allowed for the rapid and efficient placement of personnel. This advantage was a high priority for the Army, who went from 9,000 officers prior to 200,000 by the end of the War – most of whom had taken Yerkes and Terman’s test (Kevles, 1968). In total, approximately 1,750,000 individuals took the Alpha or Beta IQ test for the purpose of placement (Kevles, 1968).

The speed of Terman and Yerkes’ IQ test is attributable to a few key characteristics. The test could be given to many individuals simultaneously and was scored on an easily interpretable scale, with each soldier receiving a grade from A – E. Espeland and Sauder have highlighted the distinct advantages of such “Quantification, with its capacity to enhance predictability, coordination, and impersonality ... the hallmark of bureaucratic authority” (2007: 4). Indeed, the application of IQ tests helped create the Army’s bureaucracy (Kevles, 1968) by allowing for the impersonal assignment of responsibilities, standardization, and centralized authority (c.f. Weber, 1946). Despite these favorable characteristics, IQ testing met with considerable resistance from many of the Army’s high-ranking, pre-war officers (Kevles, 1968).

These individuals suggested that the tests had significant limitations in detecting fitness for service. It was argued that many men who performed poorly on the tests were considered

strong soldiers by their direct superiors. Further, many were unsure about the relative importance of *intelligence* for a soldier's performance in comparison to qualities such as leadership, motivation, enthusiasm, and appearance (Kevles, 1968). In addition, though Yerkes asserted that his tests measured inherent mental capacity rather than experience, men who were immigrants or illiterate frequently performed far beyond their test classification. Finally, it was claimed that highly intelligent individuals would perform poorly on the test to avoid service. These arguments were used to critique the researchers' assertions that an intelligence test was a more appropriate determinant of a soldier's placement than the opinions of seasoned officers.

Of course, the use of IQ tests also directly challenged the placement discretion of Army officers (Kevles, 1968). Not only did many resent having their authority undermined by the test, the researchers often provoked resentment by their reshuffling of recruits, particularly removing the most "promising" and depositing "below normal" men in their place (Kevles, 1968). Complaints lead to numerous investigations of the IQ-based placement system. Yerkes was especially frustrated by this, and emphasized the general consistency of test performance with supervising officers' assessments (Kevles, 1968). The persistence of IQ testing during the war, despite the resistance from many within the Army, is likely a consequence of both the pragmatic value of the tests, as well as the increased degree of openness to change that can occur during extreme events. As noted by Baron, Dobbin, and Jennings, "In periods of crisis, such as wars, stable institutions are susceptible to revolutionary change" (1986: 378).

When the war ended and the number of Army personnel decreased, a campaign by many of the career officers to end IQ testing prevailed. However, this was a temporary abeyance. Eventually, the young career officers who were selected based on their IQ scores in World War I would be influential in reinstating the IQ-based placement system for ten million draftees during World War II (Lemann, 1999: 53). Since then, even in peacetime, the Army has continued to use

IQ test scores as an important factor in placement decisions for all incoming troops (Department of Defense, 2002), and has continued to support IQ research (McHenry et al., 1990).

The Army project was important for the diffusion of IQ testing. One of the key obstacles to the adoption of an innovation is its unfamiliarity. Rogers describes this attribute of an innovation as its *observability*, that is “the degree to which the results of an innovation are visible to others” and goes on to note that “[s]ome ideas are easily observed and communicated to other people, whereas other innovations are difficult to observe or to describe to others” (2003: 258). By exposing millions of Americans to IQ testing, the Army project was an important step in making the practice of IQ-based selection more familiar, overcoming a major obstacle to diffusion. The affiliation between the Army’s IQ-based selection system and wartime success substantially altered attitudes toward the practice (Baritz, 1960: 48; Bingham, 1942). By the end of the war, IQ-based selection was familiar to a large portion of the population, and strongly associated with fast, efficient placement, and organizational success. Accordingly, public interest in the practice increased markedly after the war (Kevles, 1968).

In addition, the military project built an infrastructure for the future of IQ testing. A vast database had been created which established a distribution of IQ for a substantial portion of the male population. This database was used for post-war efforts to further validate and refine the IQ test. In the popular press, the idea of a national survey of intelligence captured the public imagination and was widely popularized in the media at the time (e.g. *Secret Mind Tests of the Army*, 1919). Further, hundreds of individuals had been trained in the science of IQ testing as test administrators and assessors. Those who were trained in and involved with the Army’s IQ testing during the war were instrumental in repurposing it for non-military uses afterwards.

IQ-reflective tests in primary and secondary schools. On the basis of Terman’s high-profile success with the Army, he received a grant from the Rockefeller Foundation to provide IQ

testing for all American school children. With help from other wartime IQ testers including Yerkes and E.L. Thorndike (Carson, 2007: 247), Terman introduced a “national intelligence test ... made available to public elementary schools” (Mathison, 2004: 4). The National Intelligence Test (NIT) was a multiple-choice exam that could be administered to groups, and was designed to put “the entire American school population on a single scale” (Carson, 2007: 247). Mathison’s historical analysis discusses how “the growth of testing in schools grew out of the ... developing emphasis on intelligence tests, particularly those that could be administered to groups rather than individuals” (2004: 4). During the 1920s, IQ tests were taken by 500,000 school children per year (Lemann, 1999: 30). By 1932, seventy-five percent of schools in large American cities had adopted the standardized IQ test (Mondale, 2001).

A number of factors contributed to the adoption of IQ testing in public schools. First, the movement toward providing universal public education in the U.S. during the beginning of the 20th century presented “the challenge of educating the new, diverse population of children [and] stimulated efforts to develop easy-to-use tests of ability to benefit instruction” (Siegler, 1992: 185). As Carson notes, by the beginning of WWII, “Intelligence as a single, measurable entity was accorded the greatest weight as a constituent of merit when there was a perceived need to sort rapidly a relatively undifferentiated mass, such as in the expanding educational systems” (2007: 233). Second, the practice of IQ testing in schools found early advocates among the progressive education movement begun a few years earlier, who championed a scientific approach to education (Mathison, 2004: 4). Third, characteristics of the test may have also fit with features of American culture. For example, Carson notes that:

“In the United States, with its highly decentralized educational system, heterogeneous population, and culture that viewed claims to expert authority skeptically, many professionals and members of the public embraced

intelligence [in the form of IQ tests] as a naturalized entity able to provide an otherwise unachievable uniformity and standardization to assessments of merit” (2007: 233).

Fourth, Carson’s historical research goes on to highlight how these applications of IQ testing were widely publicized; “Newspapers and magazines throughout the period were full of... celebrations of educational systems rearranged... admissions improved, and new prodigies discovered through testing” (2007: 255).

Finally, IQ testing provoked keen interest from the emerging educational publishing industry – e.g., companies such as Houghton Mifflin and the World Book Company (Carson, 2007: 253). Over the following years, competition among publishing companies to increase their respective share of the educational testing market lead to methodological innovations such as factor analysis, optical scanners, and matrix sampling (Madaus, Haney, & Kreitzer, 2000). Thus, after WWI, standardized IQ testing in America’s public education system had commercial advocates, popular appeal, and a large supply of Army trained psychologists available to administer tests (Carson, 2007: 254).

IQ testing for universities. Beyond primary schools, the university system was also influenced by the military’s experiment with IQ testing. Specifically, in 1926, Carl Brigham (Robert Yerkes’ assistant and a test administrator during the war) adapted the IQ test questions used by the Army to construct the Scholastic Aptitude Test for high school students (Lemann, 1999: 31). Brigham developed a conversion scale by which SAT test scores could be converted to IQ scores (Lemann, 1999: 33). Brigham’s research intended to investigate whether these IQ scores could reliably predict academic achievement. Using a sample of applicants to Princeton and Yale, as well as the military academies of West Point and Annapolis, Brigham tested the predictive validity of the SAT for college grades (Lemann, 1999: 32). While it had been

acknowledged at the time that the Army's IQ test was imperfect, it was also argued that using the test to measure university students "is still desirable ... as over 1,500,000 American men have already been rated by it" (Marvin, 1919).

The first major champion of Brigham's SAT was Harvard University President James B. Conant. Conant was interested in changing the character of the student body at Harvard. In particular, he wanted the university to draw studious, public school students from outside the private, northeastern preparatory system (Lemann, 1999: 26). At that time, however, Harvard's entrance criterion was performance on a series of essay examinations that, Conant felt, largely assessed one's familiarity with the New England prep school curriculum. Selecting students who would be successful in college from the considerably more diverse public education system was seen as necessitating a different kind of entrance exam.

Because curriculum and educational practices in America were established at the local level, Conant was adamant that students should be selected based on cognitive potential rather than knowledge or experience. When he introduced a scholarship program to Harvard for public school students, he chose the SAT test as the selection criterion explicitly because it purportedly assessed cognitive aptitude rather than achievement or experience (Lemann, 1999: 38). Thus, when applied for the first time as a selection procedure at an American university, Binet's original aim of using intelligence testing to identify potentially low performing students for additional assistance was reoriented toward the goal of selecting high performing adults for a limited number of advanced educational opportunities.

Administration of the SAT for Harvard scholarships at public high schools across the country had the effect of establishing Harvard as a national university – vastly increasing the number and diversity of applicants. The fact that students selected for the Harvard scholarships were academically successful without prep school experience provided further support for IQ-

based placement as a useful selection device (Lemann, 1999: 39). Conant's advocacy led to the introduction of similar scholarship programs at other Ivy League universities and the establishment of a national testing agency to coordinate administration of the SAT (Lemann, 1999: 39). Non-scholarship Ivy League applicants from the northeast were still admitted on the basis of their performance on essay examinations and the recommendations of preparatory schools. However, in anticipation of manpower shortages with the start of WWII, the essay entrance exams were put on what was intended to be a temporary hiatus (Lemann, 1999: 54). In its place, the much easier to administer and score SAT became the entrance exam for all applicants to the Ivy League (Lemann, 1999: 54).

After WWII, the use of the SAT as a selection criterion not only persisted, but began to diffuse outside of the Ivy League. Perhaps the most significant event supporting the adoption of SAT based selection by universities was the Serviceman's Readjustment Act of 1944. The "G.I. Bill" provided higher education tuition and living expenses for millions of returning veterans. To accommodate growing enrollment numbers, higher education went through a vast expansion. Established universities grew and new universities were founded to meet regional demands (Adams, 2000). Still, the number of applications far outpaced the expansion. For the first time, public university admissions became limited and competitive. This marked a substantial shift in American higher education. During the first quarter of the 20th century, high school students had generally received their diplomas on the basis of attendance, and universities accepted those who had a diploma, could afford tuition, and (at selective institutions) had an "appropriate" social background (Lemann, 1999: 22). When universities started to limit enrollment, many administrators looked to the SAT infrastructure developed by the Ivy League.

The factors that had made IQ testing useful for quickly processing large numbers of WWII draftees were equally useful when those returning troops enrolled in college. With the

diversity in regional education standards, SAT scores provided a measure that university admission boards could use to quickly and easily compare applicants. Advances in computing technologies developed during WWII allowed for tests such as the SAT to be graded by machines (IBM, 2008). This further lowered testing costs and scoring times (relative to alternative entrance criteria such as interviews and essay exams).

The structure of the organization that administered the SAT also played a role in the diffusion of IQ testing across universities. The Ivy League's centralized, university admissions testing agency became the non-profit Educational Testing Service in 1947, and was initially sustained by an endowment from the Carnegie Foundation (Lemann, 1999: 70-71). This fact likely contributed to the spread of the SAT in two ways. First, in order to become self-supporting, the organization needed to increase the number of universities using the SAT. Thus, the agency began a marketing campaign targeting public university administrators (Lemann, 1999: 71). Second, the Carnegie Foundation's endorsement of the initiative may have also inclined universities to adopt the SAT in the hopes of receiving educational grants from the foundation in the future (Lemann, 1999: 61). Over the course of the 1950s, the number of universities using the SAT grew to four hundred (Lemann, 1999: 85).

Early advocates of the IQ-based admissions system campaigned for the expansion of the practice. Conant published his ideas to a wide audience of educational administrators as well as the public at large. He suggested that utilizing talent efficiently, as identified by standardized IQ-reflective tests, conferred a strategic advantage in the competition against the Soviet Union for social and scientific dominance (Mondale, 2001). In addition, Conant made the case that cognitive testing was compatible with higher education's commitment to reason and "merit" (Lemann, 1999: 49).

The degree to which IQ-reflective testing's legitimacy had improved by this time is noteworthy. At universities, the adoption of IQ-based selection heralded a change in the concept of "merit" itself. Tsay, Lamont, Abbott, and Guetzkow have documented this change with their comparative analysis of letters of recommendation for the Woodrow Wilson Fellowship during the early 1950s and the late 1960s (2003). In both periods, across disciplines, recommenders' focused on the applicant's "merit". However, during the earlier period, merit was more likely to refer to an applicant's social background and moral characteristics. By the latter period, merit was substantiated by references to cognitive abilities.

In the beginning of this section, we highlighted the public's initial resistance to the notion that intelligence could be reliably measured (as expressed by Justice Goff's quote). By 1951, at the start of the Korean War, it was national policy that a college student with a sufficiently high score on an IQ test administered by ETS could defer military service (Lemann, 1999: 75). Such a policy is indicative of the extent to which *intelligence* had gained acceptance as a measurable characteristic, and the degree to which IQ testing was perceived as a legitimate basis for preferential status.

IQ testing and firms. After WWI, there was also substantial growth in the number of private organizations using standardized employment tests (Baron et al., 1986). As mentioned above, articles in national newspapers popularized the Army's success in using IQ testing for the placement of recruits (e.g. "Secret Mind Tests of the Army: Ban Lifted at Last from Censorship on Mental Exercises by Which 1,500,000 Soldiers Were Graded by Experts on Psychology for Jobs That Suited Them," 1919). The psychometric assessment techniques used in WWI were called a "war gift to industry" (Baritz, 1960: 24). Psychology became an area of interest for managers, and its selection tools became part of "standard industrial equipment" for America's increasingly large and bureaucratized corporations (48).

Many of the psychologists who had been involved in the Army's IQ testing repurposed their efforts for commercial enterprises after the war. A 1922 article in the *New York Times* announced the establishment of the Psychological Corporation for the "application of psychology to business" (*Tests to Ascertain Fitness of Workers*, 1922). The Psychological Corporation advertised its value in assisting with employee selection using "An intelligence test made ... of the general character used by the psychologists in the army to test 1,726,966 recruits" (1922).⁶

The article notes that "the stock of the corporation is owned entirely by 166 psychologists" and also identifies the corporation's executives: Lewis M. Terman, Walter Dill Scott, W.V. Bingham, Edward L. Thorndike (all Army testers and leading scholars of applied psychology), G. Stanley Hall (the first president of the APA), and James McKeen Cattell.⁷ Any profits above \$6 per share (the equivalent of \$76 in 2010) went to funding applied psychology research (Landy, 1997). Cattell, Bingham, and Thorndike in particular promoted the Psychological Corporation's tests to private firms, schools, as well as the government agencies (1997). The company would go on to be the first publisher of the Wechsler Intelligence Scale, and in 1970, was acquired by Harcourt Brace (now Pearson).

⁶ To quote the article at length: "The ordinary methods of obtaining employees are crude and wasteful, whereas in the case of typists, stenographers, clerks, bookkeepers and secretaries psychological methods of selection are now adequate. By the use of these methods perhaps \$100 on the average is saved in time and in better results each time a new employee is engaged, and there is a possibility of obtaining one worth many times that amount ... This is information that we alone have" (1922).

⁷ In addition to being the president and founder of the Psychological Corporation, Cattell was an important figure in the psychological and scientific communities of the day. He wrote the first doctoral dissertation in psychology by an American, was the first American professor of psychology, and fourth president of the APA. In addition, he served as the owner and editor of *Science* for 50 years, and founded both the *Psychological Review* and the magazine that would become *Popular Science*.

The IQ tests promoted by the psychological community offered firms a way to quickly and efficiently sort applicants, as well as identify individuals who may have sustained cognitive damage during the war (Baron et al., 1986). Firms such as Proctor & Gamble, General Electric, and Eastman Kodak hired Army trained psychologists specializing in IQ testing to assist in applicant selection and placement (Baritz, 1960). In addition to the pragmatic benefits of IQ testing for firms, the spread of the practice was also assisted by direct government interventions in personnel practices during WWII.

In their 1986 article, Baron and colleagues document the emergence of personnel departments within private organizations during this time. As militarily vital industries grew to meet wartime production demands, the federal government acted to reduce the competition for workers. This was accomplished by establishing standardized work environments, wage rates, employee classifications and promotion standards, as well as the introduction of formal job descriptions. These policies resulted in a vast increase in the prevalence of personnel departments – not only for organizations in war related industries, but also for “rival” employers. Many companies in non-critical industries were also compelled to create personnel departments in order to justify new hires and staff retention under government oversight of regional labor pools. After the war, personnel departments persisted, with one of their functions being the application of standardized hiring practices to select from the large pool of available workers (1986).

In this section, we have described the historical events that facilitated the diffusion of IQ testing to the Army, schools, universities, and firms in the United States. The predictive power of IQ, the ease with which IQ-reflective tests could be administered, and historical factors led to IQ tests becoming one of the most widely adopted products of applied psychology.

However, many nations during this period faced similar historical events and pragmatic pressures to mobilize masses of troops and universalize education *without* adopting IQ testing.

For example, despite being the birthplace of the measure, IQ tests were rarely used in France during this period. Instead, the French military system mirrored the established educational and occupational hierarchies of the day (Carson, 2007: 230). Moreover, compared to American psychologists' convergence on IQ tests as *the* measure of intelligence, the French psychological community entertained diverse conceptions of intelligence (Carson, 2007: 230). IQ tests were seen as merely one more diagnostic tool among many, and one with significantly less authority than expert judgment (231). Such comparative differences in diffusion suggest that the social structures and norms prevalent in the U.S. in the early and middle 20th century may have played a key role in the widespread adoption of IQ-reflective tests (Ferraro, Pfeffer, & Sutton, 2005).

A Focus on IQ Rather than Class or Ethnicity

The diffusion of IQ testing in the U.S. has been described as “a predictable by-product of the ongoing movement toward a meritocratic society. Progress toward a society in which people's occupations were determined by individual talents, rather than inherited position, [which] demanded some means of determining who the talented individuals were” (Siegler, 1992: 186). Interestingly, many of the resources we identify as being distributed on the basis of IQ today (i.e. education, university training, professional selection and credentials) are those that have historically been limited to individuals born into the highest socioeconomic strata of society (Bowles & Gintis, 1976; Kerckhoff, 2001). Thus, the idea that organizational resources tend to be given to those who perform well on IQ indicative tests is reminiscent of, but distinct from, Marxist critiques of how class structures may be perpetuated through resource stratification (Bowles & Gintis, 1976: Ch. 4 & 5). Indeed, a well established literature considers the influence parents' social class *still* has on important life outcomes (e.g. Kerckhoff, 1995; Nisbett, 2009;

Rothstein, 2004).⁸ However, in the following section, we focus on a dynamic that has received considerably less attention: the emergence of IQ as the dominant criterion for the allocation of resources by organizations.

Like class, ethnicity is also related to an individual's receipt of important social resources, including education (e.g. Demack, Drew, & Grimsley, 2000), networks (e.g. Dreher & Cox, 1996) and career opportunities (e.g. Pager, Western, & Bonikowski, 2009; Maume, 1999). Disparities in the receipt of these resources is one of the rationales suggested to explain correlations between ethnicity and assessments of job performance (e.g. Roth, Huffcut & Bobko,

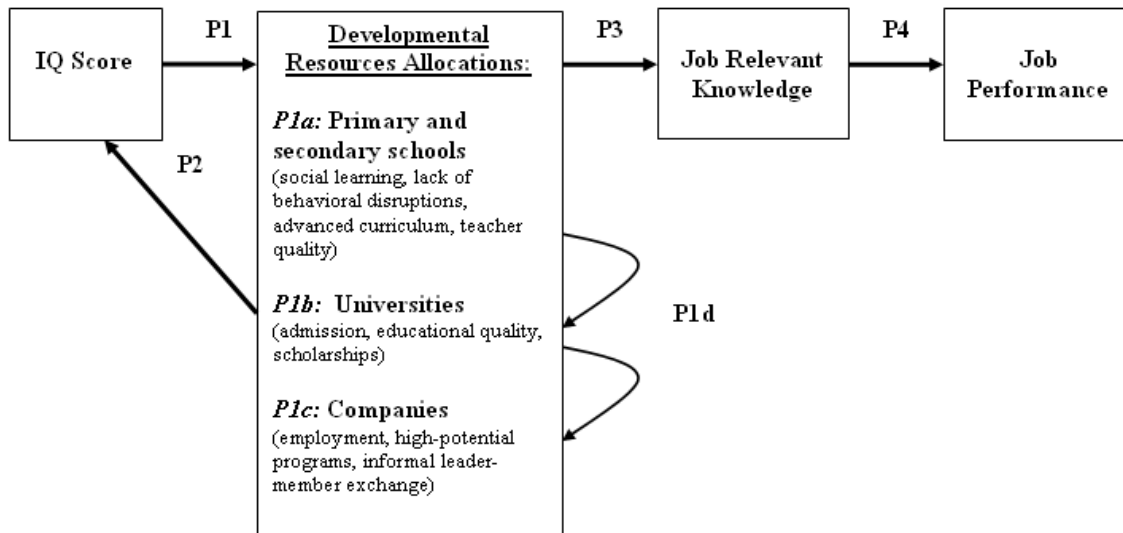
⁸ While far from definitive, a variety of recent primary and meta-analytic studies allow for direct comparison of IQ scores and measures of parents' class as predictors of important life outcomes. As compared to class background, IQ scores tend to have considerably higher correlations with four important measures of life outcomes. 1) The correlation between IQ and educational attainment (e.g. number of years of education) is estimated at .63 (Rowe, Vesterdal, & Rogers, 1998: 418); meanwhile, the correlation between parents' socioeconomic status (SES) and educational attainment is estimated at between .34 (Melby et al., 2008) and .47 (Scullin, Peters, Williams, & Ceci, 2000). 2) The correlation between IQ and occupational attainment (i.e. prestige of one's occupation) has been estimated at between .48 (Judge, Higgins, Thoresen, & Barrick, 1999) and .72 (Hunter, 1980; Jensen, 1998); meanwhile, the corrected correlation between SES background and occupational attainment is estimated at between .14 and .38 depending on sample and operationalization (Dubow, Boxer, & Huesmann, 2009; Dubow et al., 2006; Sewell, Haller, & Ohlendorf, 1970). 3) The correlation between IQ score and income is estimated at between .29 (Judge et al., 1999) and .37 (Murray, 1998; Rowe et al., 1998); meanwhile, the correlation between SES background and income is estimated at .09 (Murray, 1998). 4) Unfortunately, there are few studies of SES background and job performance, but the one we are aware of found a non-significant correlation of .06 (Colarelli, Dean, & Konstans, 1987). Of course, IQ is the strongest single predictor of job performance, estimated at between .45 and .58 (Schmidt & Hunter, 2000: 4). However, lack of reliability or variance in measures of class could be artificially suppressing effect sizes for that variable. Moreover, the above comparisons should not be interpreted as a statement that class is irrelevant in absolute terms.

2003), and rates of career advancement (Roberson & Block, 2001). While the links between ethnicity and individual outcomes are beyond the scope of this article, the arguments developed here are consistent with efforts to understand the social and organizational factors that may explain such findings. From the backdrop of the historical diffusion of the practice of IQ testing in the U.S., we now turn to the institutional factors that may inform the IQ – job performance relationship.

INSTITUTIONAL AFFECTS ON THE IQ – JOB PERFORMANCE RELATIONSHIP

In the following section, we detail how IQ scores are used by organizations as a criterion for allocating developmental resources. Further, we argue that these resources are important for acquiring competencies that facilitate future job performance. The first proposition has several sub-components, including developmental resources allocated by educational establishments and firms. By contextualizing the question of why IQ scores correlate strongly with job performance, we provide an institutional explanation for this relationship as outlined in Figure 2.

FIGURE 2
A Sociological Alternative



IQ-Centric Resource Allocation Decisions by Educational Organizations

“There are a number of reasons policy makers will continue to use test based accountability - why they find it so attractive. First, the cost and convenience of taking advantage of existing testing infrastructure and essentially refurbishing existing and well established indicators. There's the cost time-lag in using tests and comparisons to other indicators. There's the relative ease of communicating test results through various stakeholders. And finally, the power of circumventing local control by top down testing mandates. These considerations carry enormous weight and surely guarantee that test-based accountability in some form will be with us for a while” - Dr. Henry Braun (2006), Vice President of the Educational Testing Service from 1990 to 1999, on why cognitive testing in schools will continue in the foreseeable future.

In public schools, state achievement tests are commonly used as the basis for grouping students by ability level (Kerckhoff, 2001). However, a number of researchers have pointed out that achievement test scores and IQ scores are not as distinct as their labels might suggest. As Rindermann notes, “There is no theoretically important difference between intelligence tests and student assessment tests” (2008: 128). Empirical evidence also suggests that achievement tests and IQ tests are strongly related (e.g. Rindermann, 2007a⁹; Deary, Strand, Smith, & Fernandes, 2007) prompting McDaniel (2006) to argue that the “reading and math scores [on achievement tests are] a reasonable measure of state IQ” (615). Such uses of cognitive test scores to infer IQ are consistent with Rindermann’s observation that “intelligence tests and school assessment tests are alternative measures of a homogeneous construct” (2008: 129; see also Rindermann, 2007a). However, some “achievement tests” are likely to correlate more highly with IQ tests than others, and the proposition that follows holds only insofar as these “achievement tests” capture IQ.

On the basis of these tests, schools create distinct social groups (Collins, 1979; Miiller & Shavit, 1998). Ability grouping programs may include: a) creating physically separate schools that differentially prepare students for vocational or collegiate futures (i.e. “tracking”), b) maintaining the same curriculum topics within a school, but dividing students of the same age into high-middle-low classes (referred to as “XYZ classes”, “streaming”, “setting” or “banding” depending on the details of implementation), c) classes based on ability level rather than age, d) within-class groupings (e.g. reading groups based on competence level), e) enriched classes, which allow high ability students to go beyond the normal curriculum topics, and f) acceleration or deceleration programs, which allow students to finish grade levels at different speeds.

⁹ Rindermann (2007a) uses Full-Information-Maximum-Likelihood factor analysis to show that a variety of achievement tests used internationally and suggested to measure reading, math, and science competencies (i.e. PISA, TIMSS, and the PIRLS) are essentially measures of IQ.

High-quality descriptive data about the prevalence of ability grouping options has yet to be compiled by educational researchers. Complicating such efforts, programs and their implementation can differ radically at the state, city, district, school, and even classroom level. However, to the extent that ability grouping occurs in schools, there are at least four mechanisms by which students scoring higher on IQ-reflective tests may come to reside in environments with features that contribute directly to future learning and performance. These mechanisms include social learning effects, the relative absence of behavioral disruptions, advanced curriculum content, and more able teachers. These factors are considered in turn.

Premised on the principle of social learning, a significant body of research in educational sociology has shown powerful cohort affects on learning (Caldas & Bankston, 1997; Fertig, 2002; Hanushek, Kain, Markman, & Rivkin, 2003; Rindermann, 2007b; Stewart, 2008). Students clustered by ability not only learn declarative knowledge from each other (e.g. in study groups), but also come to share aspirations, motivation levels, and values (Hallilan & Williams, 1990). Social referents in the peer group to which one is assigned may contribute to the development of standards for what constitutes normal performance. Second, the prevalence of disruptive behavior in classes diminish as one moves from remedial, to regular, to advanced ability groups (Hallman & Ireson, 2005). Teacher management of such disruptions decreases the amount of class time devoted to education (2005). Further, in advanced classes, the curriculum tends to be geared toward preparing students for higher education and ultimately white collar jobs (Argys, Rees, & Brewer, 1996). Thus, the content and format of advanced courses may result in better preparation for success in college. Finally, high ability classes are more likely to be taught by teachers who have opted for, and been chosen to receive, additional education and credentials (Oakes, 1987). Collectively, these four mechanisms suggest that, due to ability grouping, individuals scoring higher on IQ-reflective tests have an increased likelihood of being taught more challenging and

socially valued knowledge by teachers with additional training and competencies in learning environments populated by more capable peers who are less disruptive. As such:

Pl1a: Primary and secondary schools tend to provide more valuable knowledge, more skilled teachers, and favorable social and disciplinary environments to students with higher IQ scores.

In *higher* education, IQ-reflective testing has been institutionalized well beyond Conant's original use of the SAT to allocate a few scholarships at Harvard (Lemann, 1999). As of 2006, 94% of American higher education institutions used the SAT in their admissions process (NACAC, 2006). Further, 59 percent of universities report that SAT scores are of "considerable importance" in making admission decisions (2006). Research by Willingham and Breland (1982) finds that the combination of grades and SAT scores explain just over two-thirds of the variance in university admission decisions in their sample. In addition, admission to graduate programs are just as contingent on IQ-reflective scores, if not more so (e.g. Landrum, Jeglum, & Cashin, 1994; Espeland & Sauder, 2007). Given the strong relationships between IQ scores and the SAT (.86 per Frey & Detterman, 2004), and IQ scores and high school grades (.50 per Neisser et al., 1996), IQ-reflective information has become *the* most important factor in university admissions (Frank, 2005; Doyle, 2006).

To give some sense of the strength of the relationship between education level achieved and IQ scores, Zagorsky (2007) analyzed the National Longitudinal Survey of Youth 1979 (N = 7403). In this sample of "Baby Boomers", the correlation between IQ scores and level of educational attainment was .62. Jensen's estimates of the relationship are even higher, at .72 (1998). If these were psychometric variables, a correlation of this magnitude would draw concerns about whether IQ scores and educational achievement were indicative of distinct constructs. Such high correlations complicate the process of disentangling the influence of the

characteristic measured by IQ scores versus the consequence of the availability of educational opportunities available to high IQ individuals.

The most selective universities are often those that provide their students with exceptional developmental resources. These may include lower pupil-teacher ratios, highly accomplished faculty, as well as access to a wide range of internships, career counseling, extra-curricular learning activities, and socialization networks (Frank, 1999, 2005). Thus, IQ-reflective information plays an important role not only in the level of an individual's educational attainment (e.g. associate degree, bachelors, masters, PhD), but also the quality of that education.

Moreover, for those high-school students who can demonstrate a high IQ through SAT scores, "merit" based scholarships allow more time to focus on their studies. These developments have led Frank to describe higher education as "the ultimate winner-take-all market" (1999, 2005).

A further indicator of the extent to which IQ testing has been institutionalized is that the rankings of universities *themselves* are partly determined on the basis of incoming classes' SAT scores. As such, university admission and scholarship decisions premised on IQ-reflective scores are part of the competition between universities for rating prestige, alumni donations, high status faculty, and corporate recruiters that offer students internships and employment opportunities. Espeland and Sauder (2007) have documented how rankings have profoundly increased the degree to which law schools weigh LSAT scores in selecting their students. The emphasis on auditable, standardized information to reflect *university* quality is compatible with, and can reinforce, the emphasis on IQ-reflective testing in admissions.

In sum, we propose that:

P1b: The quality of and opportunity for higher education is preferentially allocated to those with higher IQ scores.

IQ-Centric Resource Allocation Decisions by Employers

Companies also play a role in the distribution of scarce developmental resources. In a competitive marketplace, firms are motivated to select and invest training in individuals that they believe will be highly valuable to the firm (Lillard & Tan, 1992). In this section, we argue that firms tend to provide individuals with higher IQ-indicative credentials more opportunities to develop competencies. Specifically, we suggest that higher IQ scores increase the likelihood of a) employment, b) formal designation as a “high-potential” employee, and c) beneficial, informal leader-member exchange relationships.

Employment in general. Most basically, learning job relevant skills is assisted by actually being employed. Indeed, those with lower IQs tend to spend a larger proportion of their lives out of work (Caspi, Moffitt, Wright, & Silva, 1998), decreasing opportunities for learning valuable job skills and knowledge. One potential reason for this finding is that many employers directly or indirectly select based on IQ. In what follows, tentative evidence is compiled suggesting IQ scores play an important role in employee selection.

As reviewed earlier in this article, many scholars of industrial psychology advocate that IQ testing should be used more regularly and weighed more heavily in deciding which individuals to select and promote (e.g. Hunter & Schmidt, 1996; Schmidt & Hunter, 2000). In short, they call for greater institutionalization of the rule: select on IQ. Surveys indicate that about a third of American companies administer cognitive tests to screen applicants (Bureau of National Affairs, 1988). However, in the U.S. during the 1970s, concerns over the adverse impact of IQ-based selection on minority hiring resulted in legal rulings and public policy that increased the onus on companies to justify their use of IQ testing as a primary consideration in hiring decisions (e.g. *Griggs v. Duke Power Co.*, 1971 and EEOC regulations). The possibility of employment lawsuits against companies that openly and strongly rely on measures of IQ has

almost certainly had a negative influence on the explicit use of IQ tests by U.S. employers. However, Schmidt and Hunter note that, “Since around the mid-1980s, employers have been winning more and more such suits, and today they prevail in about 80 percent of such suits” (2000: 8). The explicit use of IQ tests for selection has been justified on the basis of the large number of studies evidencing an IQ – job performance link.

However, IQ may already play a larger role in selection decisions than acknowledged by companies. Specifically, firms frequently rely on educational credentials when recruiting and promoting employees, especially for complex jobs that require extensive training (Wilk & Cappelli, 2003; Wilk, Desmarais, & Sackett, 1995; Wilk & Sackett, 1996; Gottfredson, 1985). At the molar level, educational credentials vary in the level of attainment (e.g. secondary school, undergraduate degree, master’s degree, PhD). At the molecular level, employees can be further differentiated based on the rankings of their universities and degree programs. As we have discussed, educational level and program ranking are profoundly reflective of IQ scores. Thus, companies selecting on educational credentials may ultimately be selecting on IQ. Given the relationship between IQ scores and higher education credentials, and the professional value placed on such credentials, there is reason to suspect that candidates with high IQ scores will have comparatively better employment histories.¹⁰

“High-potential” programs. Many companies formally designate some employees as “high-potentials”, “high-flyers”, “fast-trackers”, and the like (Hall, 1976; Cox & Cooper, 1988). Programs for “high potentials” may provide an additional organizational mechanism by which those with higher IQs receive developmental resources. In a cross-European survey of

¹⁰ In addition to being employed with more regularity, high IQ individuals are selected for more prestigious jobs. Evidence of this dynamic can be found in the strong relationship between occupational prestige and the average IQ score of occupational members ($r = .90$ to $.95$; Jensen, 1998: 293).

companies, Larsen and colleagues find that a majority of sampled companies (55%) have a written or unwritten “high-flyer” policy (Larsen, London, Weinstein, & Raghuram, 1998; see Slan-Jerusalim & Hausdorf, 2007 for a similar estimate of prevalence). Larsen and colleagues define a high-flyer policy as “a systematic method to socialize, develop, and screen a group of employees who have been identified as having potential for quick promotion to top-management positions” (1998: 66). Such systematic institutional designations may provide opportunities for job rotation throughout the company, exclusive mentoring relationships, additional training programs, and special multi-source (360-degree) feedback and appraisal systems. As such, and as noted by a variety of scholars, high-flyer programs have the potential to be self-fulfilling, where the original expectation that a particular employee will become successful *leads* to their success via additional resource allocations (Hall, 1976; Larsen et al., 1998; Rosenbaum, 1979, 1984; Sheridan, Slocum, Buda, & Thompson, 1990).

A key issue is the degree to which admission into high-flyer programs is contingent on IQ-indicative information. Currently, this is an open empirical question, but there are reasons to suspect that this is often the case. For instance, many high-flyer programs recruit new university graduates who have little professional experience. Thus, selection into these programs is primarily a function of assessments of “latent potential” rather than experience or prior job performance (Bournois & Rousillon, 1992; Cox & Cooper, 1988). Proxies such as college grades and the reputation of one’s alma mater are likely to provide the initial basis for recruitment (Doherty, Viney, & Adamson, 1997).¹¹

¹¹ Of course, this is not to say that other variables (likeability, personality fit, etc.) may not also play a role in deciding who is selected as a high-flyer. Our only claim is that educational credentials (and, by derivation, IQ scores) are an important factor.

Informal leader-member exchange. A number of qualitative and quantitative studies suggest that followers perceived to be more competent are more likely to be selected for sponsorship into a special leader-member exchange relationship (Allen, 2004; Allen, Poteet, & Russell, 2000; Gerstner & Day, 1997; Kanter, 1977; Kim & Organ, 1982; Kram, 1983; Liden, Wayne, & Stilwell, 1993). These special relationships may involve mentoring and assignments to tasks that allow for learning organizationally valued skills (Dansereau, Graen, & Haga, 1975; Liden, Sparrowe, & Wayne, 1997). In short, even after individuals are selected for employment, perceptions of competence can play a role in the developmental resources they receive within the workplace.

A key question is: on what basis are such assessments of competence are made? Evidence programmatically collected by Dweck (see Dweck 2000 for a review) finds that a large proportion of Americans perceive ability in trait based terms – i.e. as an internal, stable characteristic – and, moreover, perceive intelligence as the primary driver of ability (see also Pomerantz & Saxon, 2001). Thus, managers may be particularly interested in an employees' IQ when choosing individuals to prepare for advanced responsibilities within the organization. To the extent that IQ indicative information is available in the form of educational histories and academic honors, managers may have access to a great deal of IQ-reflective information. As such, we suggest that IQ based selection occurs not only during the hiring process and during the sorting of employees into tracks, but also via managers who differentially invest time and attention into preparing promising employees for future advancement.

Above, we have described three resources that employers are likely to bestow on individuals with high IQ indicative credentials: selection, formal high potential programs, and informal leader-member exchange relationships. Further, we have discussed the extent to which

employers (particularly managers) possess, and are motivated to use IQ-indicative information to find and develop organizational members. Thus, we suggest that:

Plc: Companies tend to provide employment, formal training, and informal mentoring at a greater rate to higher IQ individuals.

Matthew Effects in Organizational Resource Allocations over Time

The “*Matthew effect*” is a reference to the biblical passage: “For unto every one that hath shall be given, and he shall have abundance: but from him that hath not shall be taken away even that which he hath” (Matthew XXV: 29). In other words, disparities in resources can be self-perpetuating over time. Merton illustrated this phenomenon with the observation that famous academics tend to receive comparatively more credit for ideas that are produced in collaboration with (or independent of, but at the same time as) less well known scientists (1968, 1988). This disproportionate credit can further increase the human, social, and cultural capital of the already successful. In economic contexts, such dynamics are usually described as increasing returns (Arthur, 1994) or as a positive feedback loop. The compounding of advantage (or disadvantage) has implications for the model developed in this article.

Specifically, Matthew effects may emerge as a consequence of the developmental resources allocated by educational institutions and employers to individuals with high IQ scores. Educational opportunities compound across grades; each year, the learning opportunities provided in the prior grade aid in the next year’s performance (and in securing future educational opportunities). For example, Stanovich (1986, 2000) reviews evidence suggesting that students who are not taught effective reading skills at a young age go on to struggle in school as reading itself becomes a vehicle for learning (see also Walberg & Tsai, 1983). The consequences of these initial disparities can last for decades (e.g. dropping out, college admission) and have been described as Matthew effects (Kerckhoff, 1993; Kerckhoff & Glennie, 1999).

As developmental resources are allocated to individuals with high IQ-reflective scores, these individuals acquire credentials and opportunities that allow them to garner additional resources. This advantage can build from grade to grade, from primary school to secondary school, from secondary school to college, from college to internships, from internships to entry positions, from entry positions to graduate school, from graduate schools to executive positions. Because of these feedback loops between resource allocations, we propose that:

P1d: Early institutional decisions to allocate developmental resources based on high IQ-reflective scores will increase the probability of further receipt of developmental resources.

A Feedback Loop Between IQ Scores and Organizational Resources.

The relative consistency of an individual's IQ scores over time has been interpreted as reflecting a stable individual difference. However, new evidence suggests that intelligence can change over time as a function of a person's life experiences (for reviews, see Flynn, 2000; Nisbett, 2009). This evidence is coincident with recent research on "neuroplasticity" which suggests that the brain is far more malleable than was previously thought (Buonomano & Merzenich, 1998). Therefore, some organizational resources allocated on the basis of IQ scores may actually result in increases in IQ scores over time. While the claim that developmental resources can influence intelligence is not theoretically novel (see Nisbett, 2009 for a review), we include a link between developmental resources and IQ scores (P2) because the potential for an individual's IQ score to change over time as a function of organizational resources has yet to be substantially incorporated into the IQ - job performance literature. In an effort to begin this discussion, we provide a brief review here.

Nisbett devotes two chapters of his recent book to detailing how differences in schooling practices may increase students' IQ (2009: Ch. 4 and Ch. 7). He reviews evidence suggesting that

smaller classes, more time in class, and teaching strategies which emphasize abstraction skills and “thinking about thinking” may all boost IQ scores. For example, with respect to the last issue, while most teaching focuses on learning specific content, an important experiment in Venezuela focused instead on teaching *how to think* (Herrnstein, Nickerson, Sanchez, & Swets, 1986). In a sample of 7th graders, the program gave lessons on classification, hypothesis testing, dimensionalization, analogizing, principles of logic, decomposing complex arguments, and evaluating the credibility and relevance of different sorts of evidence. Students who were exposed to process-based instruction performed .43 standard deviations better on an IQ test as compared to those provided with the regular, content-based instruction. Indeed, Williams (1998) and Blair and colleagues (Blair, Gamson, Thorne, & Baker, 2005) each argue that one reason that IQ scores may be increasing across generations (see Flynn, 2000) is that classroom instruction is increasingly emphasizing abstract, perceptual tasks and how to learn. Such instructional strategies, as well as the provision of smaller classes, more time on task, and talented teachers, are developmental resources likely to vary across classrooms, schools, and nations.

This raises the possibility of “Matthew effects” between current and future IQ scores. For example, Shaywitz and colleagues (1995) have shown that a child with an IQ of 120 in first-grade will tend to have an IQ of 135 by fifth-grade; conversely, a child with an IQ of 80 in first grade will tend to have an IQ of 74 by fifth grade. Thus, in this sample, IQ scores “fanned out” overtime, an indication of Matthew effects. Empirically modeling such recursive effects would require longitudinal data on the variables involved (e.g. IQ scores, developmental resource allocations, etc.) in conjunction with sophisticated analytic techniques (e.g. time series analysis and Granger-causality equations; see Granger & Newbold, 1986). While acquiring such data

would be challenging, these recursive relationships may end up playing a critical role in explaining the institutionalized link between IQ scores and job performance.¹²

P2: Allocation of organizational resources will increase IQ scores.

Developmental Resources, Job Relevant Knowledge, and Job Performance

The key argument of Proposition 3 is very straightforward. People who receive a higher quality education, more educational opportunities, scholarships, and exceptional career training, will develop job-relevant knowledge. Those with higher IQ scores will have an increased likelihood of receiving the developmental resources described in Proposition 1. Therefore, higher-IQ scores will lead to greater job relevant knowledge through the mediating mechanism of the institutional allocation of developmental resources. Proposition 4 suggests that such opportunities to develop job relevant knowledge will tend to result in better performance. While these propositions may seem obvious, obvious does not imply tautological. Learning opportunities are not always turned into useful knowledge. Greater capabilities do not necessarily lead to better job performance (for example, when motivation is very low).

P3: Superior developmental resources result in the acquisition of more job relevant knowledge.

P4: More job relevant knowledge leads to improved job performance.

¹² In an important article, Dickens and Flynn (2001) mathematically formalize such feedback loop models, and apply them to the question of whether IQ is environmentally or genetically determined. Their conclusion is that individuals with higher IQs come to reside in IQ-supportive environments that then aid in further developing their cognitive capacities – i.e. that IQ is best thought of as a product of recursive interactions between nature and nurture. Although Dickens and Flynn emphasize self-selection into cognitively rich environments (e.g. higher IQ individuals choose to read more and thereby increase their intelligence), the institutionalized practice of allocating scarce developmental resources based on IQ scores would accelerate such recursive dynamics.

In many ways, these predictions are consistent with the existing literature. Prior research suggests that individuals with higher IQ scores perform better because they acquire more job-related knowledge (e.g. Hunter, 1986). However, as noted earlier, prior research tends to assume that job-related knowledge is learned exclusively at the job in question – an assumption we reject.

SELF-FULFILLING DYNAMICS

Social science hopes for its products to be understood and applied by policy makers and practitioners. Paradoxically, the realization of this hope makes social science uniquely susceptible to the problem of the “double hermeneutic” (Giddens, 1987). A hermeneutic is a process of interpretation, and Giddens’ notion of a double hermeneutic describes recursive interactions between social science findings and subject behavior. As described by Giddens, these interactions occur because “the concepts of the social sciences are not produced about an independently constituted subject-matter, which continues regardless of what these concepts are. The ‘findings’ of the social sciences very often enter constitutively into the world they describe” (1987: 20). Such recursive dynamics have also been explored in terms of “performativity” (Callon, 1998; MacKenzie, 2004; MacKenzie, Muniesa, & Siu, 2007). These scholars have argued that the relationships described by social science end up becoming part of reality as actors attempt to “perform” in ways consistent with these relationships.

Our propositions suggest that the IQ – job performance relationship is partly self-fulfilling, where institutional architects (e.g. social scientists, military leaders, educators, and managers) have adopted the belief that IQ determines performance, employed IQ testing as a basis for the allocation of scarce developmental resources, and collectively caused their beliefs to become more true due to their actions. Thus, this article follows in the footsteps of scholars who have described how the institutionalization of assumptions from economics, most notably the norm of self-interest, tends to reproduce behavior consistent with those assumptions (Ferraro et

al., 2005; Ghoshal, 2005; Ghoshal & Moran, 1996; Miller, 1999; Ratner & Miller, 2001). In their article on the self-fulfilling nature of economic assumptions, Ferraro and colleagues suggest that future research should “look for the possibility of self-fulfilling feedback loops in theories of behavior other than economics” (2005: 20), and mention assumptions about intelligence as a promising area for such an inquiry. We agree and have outlined how the organizational applications of IQ scores as a basis for allocation of developmental resources may reinforce the IQ - job performance finding.

However, our use of term *self-fulfilling* to describe the IQ - job performance link should not be misinterpreted as implying that IQ would have *no* relationship with job performance absent its institutionalization as a basis for resource allocation. Rather, if forced to speculate, we would suggest that a weak-moderate relationship between IQ and job performance could be transformed into a strong relationship due to the institutionalization of IQ testing as a basis for resource allocation. Moreover, it should also be noted that our theoretical model does *not* imply that IQ scores are a useless criterion for managers to include in selection decisions. As our historical discussion suggests, the general usefulness of IQ testing in partially predicting performance has contributed to its diffusion. However, once the idea that higher IQ individuals perform better becomes a part of widespread organizational resource allocations, the “pure” influence of IQ becomes confounded with the opportunities bestowed on the basis of that characteristic. Thus, the application of one of industrial psychology’s most promising findings highlights the difficulty inherent in studying a population where social science observations can become institutionalized facts.

DISCUSSION

To summarize our arguments, we suggest that institutionalized practices enable individuals with high IQ scores to develop better capabilities. As a consequence, the observed relationship between IQ scores and job performance is confounded by the fact that opportunities have been allocated based on IQ-reflective criteria. We have argued that the cumulative result of these developmental resource allocations by educational establishments and companies is a stronger empirical relationship between IQ scores and job performance.

Moreover, in our analysis of the diffusion of IQ-testing, we have suggested that this differential allocation is a historically recent phenomenon (starting in the early part of the 20th century). If developmental resources were still primarily allocated on the basis of social background, then we would expect a weaker relationship between IQ scores and job performance (and a stronger relationship between social background and job performance). The same logic would apply for other individual differences on which performance enhancing resource allocations might be premised (e.g. physical attractiveness, ethnicity, gender, or conscientiousness). While we are very aware of the predictive advantages of applying standardized cognitive selection, we suggest that the prescription to “select on intelligence” may need to be bounded by institutional context. The remainder of the article is devoted to the managerial and public policy implications of our theory, as well as its relevance for future research in organizational behavior.

Potential Managerial Implications

Given the potential for a self-fulfilling dynamic, the managerial implications for selecting employees based on IQ indicative credentials are subtle. If the IQ – job performance relationship is (at least partly) institutionally created, how should managers and HR professionals respond to the oft-repeated prescription “select on intelligence”? In our view, the answer to this question hinges on the institutional context in which managers find themselves. In the United States, IQ-reflective criteria may continue to be a useful basis for selection *because* high IQ individuals differentially possess institutionally endowed tools and competencies. Moreover, in professional, knowledge intensive industries such as consulting, employees’ IQ-reflective credentials may enhance organizational legitimacy and prestige. Therefore, adopting socially legitimate signifiers of competence (i.e. employing individuals with prestigious credentials) may be of organizational value in and of itself (Heugens & Lander, 2009).

On the other hand, the Western manager who attempts to apply the “select on intelligence” prescription in countries with very different institutional designs and histories may have less success. China offers an illustrative example of how a different institutional context might undermine the managerial edict “select on intelligence”. During the 1960's and 1970's, in association with China's “cultural revolution”, access to educational resources were greatly reduced for everyone, and the education that did occur was distributed in a largely egalitarian and undifferentiated fashion (Yang R., 2004). In addition, Chinese students tend to believe that motivation, rather than general cognitive ability scores, predicts successful task accomplishment (Grant & Dweck, 2001, Yang S-Y. & Sternberg, 1997). Taking these institutional and cultural facts together, one would expect a weaker IQ score – job performance relationships in China than in the United States. As such, a manager in China (perhaps from an American-based multinational company) who “selects on intelligence” may overlook more competent employees. In this context, selection on the basis of characteristics by which resources are locally allocated (e.g. guanxi or social connections) may prove beneficial.

Currently, the above suggestion is purely speculative. Outside of developed Western countries, Australia, and Japan, there is a vacuum of evidence to validate (or refute) the claim that IQ predicts job performance. As noted by Sternberg, Grigorenko, and Bundy, “IQ has been most studied where it was invented and where it is most appreciated, that is, in the established market economies and especially in the United States” (2001: 2). Below, we offer a brief sketch of how the allocation of institutionalized resources may vary across cultures.

Cross-country comparison of the allocation of educational resources based on IQ.

While IQ testing in educational contexts first took hold in America, it has since spread to other countries. In particular, throughout Northern Europe (e.g., United Kingdom, Belgium, Netherlands, Denmark, Norway, Sweden, Finland, Germany), IQ-reflective tests play a role in

sorting students into tracks (Kerckhoff, 2001; Miiller & Shavit, 1998; Shavit & Blossfeld, 1993). In these countries, a common practice is to create physically separate schools based on rankings of student ability, which are derived from some combination of IQ-reflective tests and teacher recommendations. For example, in the Netherlands, at age 12, students are sorted into an HBO track (vocational training) or an MBO track (university preparation) based on test scores, with little movement between tracks. Due to their colonial association with Britain, Jamaica has a similar educational structure, with Grade Six Achievements tests (GSAT) determining whether students go into an “upper track” (college bound) or a “lower track”. Turner (1960) calls these sorts of educational structures “sponsored mobility systems”. In contrast, primary and secondary school in the United States have what Turner (1960) calls a “contest mobility system”, which allows for yearly movement between basic and enriched curricula.

Kerckhoff finds strong evidence of stratification and accumulative advantage in both Britain (1993) and in the United States (Kerckhoff & Glennie, 1999), but suggests that these systems are so different that it is difficult to directly compare the strengths of these effects. Further complicating matters, many countries cannot be easily classified as a contest or sponsored mobility system. For example, the Russian curriculum is standardized until age 15 (i.e. no mobility), then has clear vocational and university preparatory tracts (i.e. sponsored mobility), then uses a test modeled on the SAT as the basis for competitive admissions to universities which differ widely in quality (i.e. a contest mobility system). In many developing countries, there is no national system, and educational practices have been decentralized to the local level (Arnové & Clements, 2004).

In sum, nations vary in the degree to which they distribute educational resources on the basis of IQ-reflective scores. Further, different kinds of ability based groupings occur across and within countries, and these groupings occur at different ages. Thus, we would expect that P1a-d

will be stronger in some regions and weaker in others. While it is beyond the scope of this paper to compare in detail the international educational systems of various nations (but see Kerckhoff, 2001; Miiller & Shavit, 1998; Shavit & Blossfeld, 1993), in our discussion of future research opportunities, we return to the value of investigating between nation variance in IQ-based resource stratification.

Cross-country comparison of IQ-based resource allocations in the workplace. In addition to educational practices, the prevalence of IQ testing in employment screening also varies substantially across countries (for a recent review, see Lievens, 2007). In one of the more comprehensive studies to date, Ryan, McFarland, Baron, and Page (1999) survey companies in 20 countries about their use of cognitive tests. They find that use of such tests range from “almost always” (80-100%: e.g., Belgium) to “rarely” (0-20%: e.g., Italy); in particular, to rank countries from most to least frequently used: Belgium, Netherlands, Spain, New Zealand, Portugal, South Africa, United Kingdom, Sweden, Singapore, Ireland, Canada, Greece, Australia, France, United States, Germany, Hong Kong, Italy, Japan, and Malaysia. Interestingly, the United States comes in at the low end of the list, ranking closest to the option “occasionally: 21-50%”. This is consistent with other studies showing that about a third of American companies use cognitive ability tests (Bureau of National Affairs, 1988). Such evidence suggests that some western nations may have embedded IQ testing into institutional designs to an even greater degree than the United States. Indeed, this is one potential explanation for Salgado and colleagues’ (2003a) finding that the IQ – job performance relationship is stronger in their (largely northern) European meta-analysis than in other meta-analyses from North America.

The arrow of history. We have noted that nations vary in their institutionalization of IQ-testing as a basis for the allocation of developmental opportunities. One force that supports greater institutionalization of IQ-based testing in resource allocation is the adoption of Western

education practices abroad. For example, open borders and lower transportation costs have increased the international movement of college students (Bauder, 2006). As such, many universities have “gone global”, and seek to draw students from around the world. However, selecting internationally is complicated by diverse educational curricula and evaluation systems. IQ-based selection, as modeled by the American university system, provides a way in which students from diverse educational backgrounds can be placed on a common continuum. Further, adopting selection criteria such as the SAT, GRE, MCAT, and GMAT also allows non-U.S. universities to be included in (usually western) university rankings in the competition for global prestige, recruiter attention, and prospective students. For western organizations expanding abroad, exporting familiar, IQ-based selection practices may provide an expedient way to assess individuals with unfamiliar backgrounds and work experiences. As such, we suspect that the “arrow of history” points in the direction of greater use of IQ scores as a basis for resource allocation.

In short order, we propose a research agenda for assessing the impact of these trends and the potential for validating or repudiating our theory. In addition, we discuss how this theory fits into the larger project of appreciating how institutional designs influence organizational behavior theories. But first, the potential public policy implications of the theory are discussed.

IQ-centric Resource Allocations & Public Policy

“If the intelligence test really measured the unchangeable hereditary capacity of human beings, as so many assert, it would inevitably evolve from administrative convenience into a basis for hereditary caste” (Lippmann, 1922: 297).

There are public policy choices associated with how to allocate many of the developmental resources discussed in this article. In educational policy and case law, there have been heated debates regarding the acceptability of IQ - reflective selection practices. To a great

extent, there are two distinct camps associated with these resource allocation debates: those who support stratification in developmental resources based on estimates of ability, and those who believe that developmental resources should be allocated relatively equally, irrespective of ability differences. The first position, described here as “pro-merit”, supports proposals to provide differential educational programs (e.g. tracking, streaming, banding) within and across schools, and “merit based” competition for college admissions and scholarships (e.g. IQ-reflective scores and secondary school grades) (Gottfredson, 1986, 1988; Herrnstein & Murray, 1994). Such criteria are also recommended for workplace selection, a position epitomized by Hunter and Schmidt:

“What should social policy be in the area of employee selection? Social policy should encourage employers to hire on the basis of mental ability. ... [S]uch a policy is likely to maximize economic efficiency and growth (including job growth), resulting in increases in the general standard of living” (1996: 467).

The second position (described here as “pro-equality”) supports a very different set of public policies, such as: “heterogeneous ability grouping” and “inclusive” primary and secondary education. Under these policies, students of varying ability levels (e.g. “gifted” and “disabled”) are taught together, university scholarships and admissions emphasize “need”, and employers’ selection and promotion practices should minimize “adverse impact”. In the domain of educational policy, Mensh and Mensh epitomize this position.

“[Schools] have to reverse their priorities so that education, not assessment, becomes the primary objective, and at the same time [need to] strive to overcome racial, class, and sex discrimination. [Schools should] encourage all children to aspire and provide them with equal opportunities to learn” (1991: 158).

To some extent, these different sets of public policy recommendations derive from contrasting ideologies and values. However, these perspectives also emerge from divergent understandings of empirical realities. For example, a classic empirical question is the extent to which IQ scores are changeable as a function of “the environment” (c.f. Nisbett, 2009; Jensen, 1998). The “pro-merit” camp suggests that differences in ability are mostly immutable (Herrnstein & Murray, 1994). To quote Gottfredson:

“[A] society’s collective welfare requires that it make good use of its most talented members but also find productive roles for its least talented. ... Much social policy has long been based on the false presumption that there exist no stubborn or consequential differences in mental capability. Worse than merely fruitless, such policy has produced one predictable failure and side effect after another, breeding widespread cynicism and recrimination” (1997: 123-124).

Meanwhile, the “pro-equality” position suggests that the practices of parents, teachers, principals, professors, and community leaders make an enormous difference in both intelligence and general competence, and that this empirical reality demands that resources should be stratified as little as possible (Kamin, 1974; Nisbett, 2009; Flynn, 1999). Evidence that IQ -reflective tests appear to strongly correlate with variables that have historically been the basis of social divisions (e.g. socioeconomic status and ethnicity) also provokes concern as to the legitimacy of IQ as a basis for differentiation. For example, research by Rothstein (2004) finds that ethnicity and socioeconomic background are associated with SAT scores, and that the relationship between SAT scores and college performance is dramatically weakened after statistically controlling for these background characteristics.

For those concerned with distributive justice (e.g. Rawls, 2001), the potentially self-fulfilling dynamics associated with the IQ – job performance relationship challenge the standards

of distributive equality. Institutionalized resource allocation practices that favor those who are already advantaged in life. Intelligence testing was originally envisioned by leaders like James Conant as a tool for allocating developmental resources on a criterion more meritorious than social class. However, to the extent that IQ becomes a basis for differentiating developmental opportunities, it could be argued that one form of caste system will have merely replaced others. Indeed, in his classic dystopia, Young (1958) suggested that placing students on career tracks based on IQ scores recreated many of the traditional bases of social divisions from previous eras (e.g. heredity, socio-economic status, social networks). In this way, research such as that by Rothstein raises questions about what IQ-reflective tests measure and the appropriateness of the characteristic as a determinant of organizational resource allocations. While our model (as summarized in Figure 2) is amenable to different interpretations of what IQ tests measure, this question has implications for public policy. If IQ tests measure innate potential – as generally assumed by Schmidt and Hunter – then ours is a story of innate potential leading to superior developmental opportunities, leading to job relevant knowledge, leading to job performance. If, as argued by Sternberg and Nisbett, IQ tests measure both innate potential and prior learning, then ours is a story of some indeterminate combination of potential and knowledge leading to superior developmental opportunities, leading to job relevant knowledge, leading to job performance. If, as Mensh and Mensh suggest, IQ tests are measures of knowledge, then ours is a story of knowledge disparities leading to institutionally allocated opportunities to acquire more knowledge, and culminating in better job performance.

This article adds a further empirical dimension to the debate over policy alternatives. Specifically, to the extent that the link between IQ and ultimate job-performance is mediated by developmental resources allocated on the basis of IQ-reflective scores, the case for allocating developmental resources based on IQ is weakened. Said another way, if a society could allocate

resources on a given criterion (say, eye-color) and produce a strong correlation with job performance becomes *less* justifiable. Conversely, the IQ-based allocation of developmental resources becomes *more* justifiable to the extent that: 1) variance in ultimate job performance is explained by immutable differences in mental processing and 2) these differences are captured by IQ tests. In short, there are a number of empirical questions that need to be answered before it is possible to make moral or policy judgments. Foremost: to what extent is the connection between IQ and job performance a function of institutionalized, IQ-based allocations of developmental resources? In next section, we sketch some possibilities for discovering the answer to this question.

Future Research

We have presented empirically grounded reasons to suspect a self-fulfilling dynamic given extant institutionalized practices, but not direct empirical testing of such effects. Historical analysis offers only an initial step in substantiating these arguments. Rigorously test these ideas entails showing that the relationship between IQ and job performance is positively associated with the degree to which IQ-reflective scores are an institutionalized basis for resource allocations. For example, meta-analytic techniques could provide evidence as to whether the strength of the IQ – job performance relationship is moderated by the degree of IQ-based ability grouping in education (e.g. “tracking” versus “inclusive” educational designs).

Perhaps the strongest evidence would derive from “natural experiments”. Researchers could investigate the relationship between IQ and job performance for workers of different ages in regions in which there have been major shifts towards or away from stratification of educational or workplace resources based on IQ. To illustrate, we again turn to China. Due to Deng Xiaoping's widespread educational reforms since the 1980s, access to educational resources have become increasingly stratified based on intellectual ability (Yang R., 2004). Thus, the

strength of the IQ score – job performance relationship should become stronger over time for Chinese workers. One pragmatic way to test this would be to compare the strengths of IQ score – job performance correlations for employees who were educated pre-1980s versus post-1980s, but who hold the same jobs. If the correlations are stronger for post-1980s employees, then this would provide some support for the IQ – job performance relationship as a function of stratified IQ-based resource allocations. Beyond China, future research on the influence of resource allocations could build on the literature comparing educational (Kerckhoff, 2001, Shavit & Blossfeld, 1993) and workplace (Lievens, 2007) practices across countries. Of course, such natural experiments also occur within North America (at the state / province level), but the variation in institutional design may be considerably smaller than between nations.

Both meta-analytic moderation and natural experiments are made more robust to the extent that they rule out a number of confounds. For example, given that IQ score is expected to be more predictive of performance in complex tasks (Kuncel et al., 2004), controlling for differences in the complexity of the economic activities engaged in by different populations is critical. Other controls include range restriction in a populace's IQ scores and the degree to which there is an assortive matching process by which people possessing high IQs tend to be located in more complex jobs (Wilk et al., 1995). In addition, causal claims may be enhanced with analytic techniques such as time series analysis and Granger-causality equations (Granger & Newbold, 1986). Collectively, such empirical tests can allow us to assess the existence and effect sizes of the institutional production of the IQ – job performance relationship. Answering these questions offers a number of exciting opportunities for future research.

Broader Implications for Organizational Behavior Research

Within in the field of organizational behavior, there are often-repeated requests for more contextualization of our research (Brockner, 2003; Cappelli & Sherer, 1991; Rousseau & Fried,

2001; Johns, 2006). The most common version of such a concern is that culturally constructed beliefs and meanings moderate our ontological claims – e.g. that the findings from Western samples cannot be automatically generalized to other cultures. In recent years, and as more non-Westerners join the academy, the community of organizational behavior researchers involved in cross-cultural research has grown (Tsui, Nifadkar, & Ou, 2007). However, while this work is incredibly important, much of this research is oriented toward assessing variations in *beliefs* (e.g. endorsement of individualistic versus collectivistic orientations) across countries. Meanwhile, the study of variations in *institutional* designs may also serve to specify the boundaries of western organizational behavior findings, as well as reveal which institutional practices are relevant for organizational behavior in particular contexts.

Although individual beliefs and institutional designs often pull in the same direction, this is not always the case. For example, an important program of research by Yamagishi and colleagues (Yamagishi, 1988, Yamagishi, Cook, & Watabe, 1998; Yamagishi & Yamagishi, 1994) contradicts the stereotype that Japanese culture engenders greater trust among individuals than American culture. Instead, the high levels of cooperation observed in Japan appear to be a consequence the strict formal and informal institutional sanctions that promote interdependence; however, the individual propensity to trust others is actually relatively low (see also Huff & Kelley, 2003; Hagen & Choe, 1998). In short, it can be misleading to focus on psychological variation while ignoring institutional variation. Our article is part of the effort to correct this tendency by considering how different institutional practices also serve to bound our ontological claims about organizational behavior. Specifically, we suggest that the relationship between IQ and job performance is strongest in contexts where institutional architects have *already* adopted this edict, and in so doing, helped create its validity.

Conclusion

Social science findings that become the basis for organizational policies can have unintended consequences. The strongest findings are particularly susceptible to the problems described here as they are more likely to be widely adopted in organizational practices. We have suggested that IQ scores have become embedded in organizational resource allocation decisions in ways that can have self-fulfilling consequences. This article complicates the story researchers have been telling about the relationship between IQ and job performance. While not denying that intelligence is a critical resource in performing one's job, we have provided reasons to suspect that institutionalized practices outside the workplace have the potential to inflate the IQ score – job performance relationship in ways that must also be recognized in our theories and empirical research. The aim is not to attack one of the more compelling findings from organizational scholarship, but instead to supplement it with a measure of disciplined skepticism.

REFERENCES

- Adams, J. A. (2000). The G.I. Bill and the changing place of U.S. higher education after World War II. *Annual Meeting of the Association for the Studies of Higher Education*. Sacramento, CA.
- Allen, T. D. (2004). Protégé selection by mentors: Contributing individual and organizational factors. *Journal of Vocational Behavior*, 65, 469-483.
- Allen, T. D., Poteet, M. L., & Russell, J. E. A. (2000). Protégé selection by mentors: What makes the difference? *Journal of Organizational Behavior*, 21, 271-290.
- Argys, L. M., Rees, D. I., and Brewer, D. J. (1996). Detracking America's schools: Equity at zero cost? *Journal of Policy Analysis and Management*, 15, 623-645.
- Arnone, R. F., & Clements, M. M. (2004). Education and international issues of social mobility of underprivileged groups. In J.W. Guthrie (Ed.), *Encyclopedia of Education*.
- Arthur, W. B. (1994). *Increasing returns and path dependence in the economy*. Ann Arbor: University of Michigan Press.
- Baritz, L. (1960). *Servants of power*. Wesleyan University Press, Middletown, CT.
- Baron, J. N., Dobbin, F. R., & Jennings, P. D. (1986). War and peace: The evolution of modern personnel administration in U. S. industry. *American Journal of Sociology*, 92, 350-383.
- Bauder, H. (2006). *Labor movement: How migration regulates labor markets*. New York: Oxford University Press.
- Becker, G. S. (1993). *Human capital: A theoretical analysis with special reference to education* (3rd ed.). New York: Columbia University.
- Bingham, W. V. (1942). How the Army sorts its manpower. *Harper's Magazine*, November: 185.
- Blair, C., Gamson, D., Thorne, S., & Baker, D. (2005). Raising mean IQ: Cognitive demand of mathematics education for young children, population exposure to formal schooling, and the neurobiology of the prefrontal cortex. *Intelligence*, 33(1), 93-106.
- Blundell, R., Dearden, L., Meghir, C., & Sianesi, B. (1999). Human capital investments: The returns from education and training to the individual, the firm and the economy. *Fiscal Studies*, 20(1), 1-23.
- Bourdieu, P. (1986). The forms of capital. In J. G. Richardson (Ed.), *Handbook of theory and research for the sociology of education* (pp. 241-258). New York: Greenwood.
- Bournois, F. & Roussillon, S. (1992). The management of high flyer executives in France: The weight of national culture. *Human Resource Management Journal*, 3(1), 37 – 56.
- Bowles, S., & Gintis, H. (1976). *Schooling in capitalist America*. New York: Basic Books.
- Braun, H. (2006, October 16). Boisi Lecture: Testing and Educational Policy. Retrieved April 9, 2008, from <http://at.bc.edu/podcasts/podcast.xml>
- Brockner, J. (2003). Unpacking country effects: On the need to operationalize the psychological determinants of cross-national differences. *Research in Organizational Behavior*, 25, 333-367.
- Buonomano, D. V., & Merzenich, M. M. (1998). Cortical plasticity: from synapses to maps. *Annual Review of Neuroscience*, 21, 149-186.
- Bureau of National Affairs. (1988). Recruiting and selection procedures. *Personnel Policies Forum Survey*, No. 146. Author, Washington, DC.
- Caldas, S. J., & Bankston, C. (1997). Effect of school population socioeconomic status on individual academic achievement. *Journal of Educational Research*, 90, 269-277.
- Callon, M. (1998). Introduction: The embeddedness of economic markets in economics. In M. Callon, (Ed.), *The laws of the markets*. Oxford, UK: Blackwell.

- Campbell, J. P., Gasser, M., & Oswald, F. (1996). The substantive nature of job performance variability. In K. Murphy (Ed.), *Individual differences and behavior in organizations*. Hillsdale, NJ: Erlbaum
- Cappelli, P., & Sherer, P. D. (1991). The missing role of context in OB: The need for a meso-level approach. *Research in Organizational Behavior*, 13, 55-110.
- Carson, J. (2007). *The measure of merit: Talents, intelligence, and inequality in the French and American republics, 1750-1940*. Princeton, NJ: Princeton University Press.
- Caspi, A., Moffitt, T. E., Wright, B. R. E., & Silva, P.A. (1998). Early failure in the labor market: childhood and adolescent predictors of unemployment in the transition to adulthood. *American Sociological Review*, 63, 424-51.
- Colarelli, S. M., Dean, R. A., & Konstans, C. (1987). Comparative effects of personal and situational influences on job outcomes of new professionals. *Journal of Applied Psychology*, 72, 558-566.
- Collins, R. (1979). *The credential society: A historical sociology of education and stratification*. New York: Academic Press.
- Cote, S., & Miners, C. T. H. (2006). Emotional intelligence, cognitive intelligence, and job performance. *Administrative Science Quarterly*, 51(1), 1-28.
- Cox, C. J., & Cooper, C. L. (1988). *High flyers: An anatomy of managerial success*. New York: Blackwell.
- Dansereau, F., Graen, G., & Haga, W. (1975). A vertical dyad approach to leadership within formal organizations. *Organizational Behavior and Human Performance*, 13, 46-78.
- Deary, I.J., Strand, S., Smith, P., & Fernandez, C. (2007). Intelligence and educational achievement. *Intelligence*, 35, 13-21.
- Demack, S., Drew, D., & Grimsley, M. (2000). Minding the gap: ethnic, gender and social class differences in attainment at 16, 1988 - 95. *Race, Ethnicity and Education*, 3, (2), 117 - 143.
- Department of Defense. (2002). The recruiting process. Population representation in the military services. http://www.pentagon.mil/prhome/poprep2002/chapter2/c2_recruiting.htm. Retrieved Mar. 9, 2008.
- Dickens, W. T., & Flynn, J. R. (2001). Heritability estimates versus large environmental effects: The IQ paradox resolved. *Psychology Review*, 108, 346-369.
- Dilchert, S., Ones, D. S., Davis, R. D., Rostow, C. D. (2007). Cognitive ability predicts objectively measured counterproductive work behaviors. *Journal of Applied Psychology*, 92, 616-627.
- Doherty, N., Viney, C. and Adamson, S.J., (1997). Rhetoric or reality: shifts in graduate career management? *Career Development International*, 2, 173-179.
- Doyle, W. R. (2006). Adoption of merit-based student grant programs: An event history analysis. *Educational Evaluation and Policy Analysis*, 28, 259-285.
- Dreher, G. F., & Cox, T. H. (1996). Race, gender, and opportunity: A study of compensation attainment and the establishment of mentoring relationships. *Journal of Applied Psychology*, 81, 297-308.
- Dubow, E. F., Boxer, P., Huesmann, L. R. (2009). Long-term effects of parents' education on children's educational and occupational success: Mediation by family interactions, child aggression, and teenage aspirations. *Merrill-Palmer Quarterly*, 55, 224-249.
- Dubow, E. F., Huesmann, L. R., Boxer, P., Pulkkinen, L., & Kokko, K. (2006). Middle childhood and adolescent contextual and personal predictors of adult educational and occupational

- outcomes: A mediational model in two countries. *Developmental Psychology*, 42 (5), 937-949.
- Dweck, C. S. (2000). The development of ability conceptions. In A. Wigfield, J. S. Eccles (Eds.), *Development of achievement motivation* (pp. 57-88). San Diego, CA: Academic Press.
- Espeland, W. N., & Sauder, M. (2007). Rankings and reactivity: How public measures recreate social worlds. *American Journal of Sociology*, 113(1), 1-40.
- Ferraro, F., Pfeffer, J., & Sutton, R. I. (2005). Economics language and assumptions: How theories can become self-fulfilling. *Academy of Management Review*, 30(1), 8-24.
- Fertig, M. (2002). Educational production, endogenous peer group formation and class composition. Evidence from the PISA 2000 study. Essen: RWI Discussion Papers, 2.
- Flynn, J. R. (1999). Searching for justice: The discovery of IQ gains over time. *American Psychologist*, 54, 5-20.
- Flynn, J. R. (2000). IQ trends over time: Intelligence, race, and meritocracy. In K. Arrow, S. Bowles, & S. Durlauf (Eds.), *Meritocracy and economic inequality*. Princeton, NJ: Princeton University Press.
- Frank, R. H. (1999). Higher education: The ultimate winner-take-all market? (CHERI Working Paper #2). Retrieved February 9th 2008, from Cornell University, ILR School site: <http://digitalcommons.ilr.cornell.edu/cheri/2>.
- Frank, R. H. (April 14, 2005). Intense competition for top students is threatening financial aid based on need. *The New York Times*, Retrieved February 1, 2008, from <http://www.nytimes.com/2005/04/14/business/14scene.html?oref=login>.
- Frey, M. C., & Detterman, D.K. (2004). Scholastic assessment or g? The relationship between the scholastic assessment test and general cognitive ability. *Psychological Science*, 15, 373-378.
- Ganzach, Y. (1998). Intelligence and job satisfaction. *Academy of Management Journal*, 41, 526-539.
- Gardner, H. (1999). *Intelligence reframed: Multiple intelligences for the 21st century*. New York: Basic Books.
- Gerstner, C. R., & Day, D. V. (1997). Meta-Analytic review of leader-member exchange theory: Correlates and construct issues. *Journal of Applied Psychology*, 82, 827-844.
- Ghoshal, S. (2005). Bad management theories are destroying good management practices. *Academy of Management Learning and Education*, 4, 75-91.
- Ghoshal, S., & Moran, P. (1996). Bad for practice: A critique of the transaction cost theory. *Academy of Management Review*, 21, 13-47.
- Giddens, A. (1987). *Social theory and modern sociology*. Cambridge, UK: Polity Press.
- Gladwell, M. (2008). *Outliers: The story of success*. London, UK: Little, Brown, & Co.
- Gottfredson, L. S. (1985). Education as a valid but fallible signal of worker quality: Reorienting an old debate about the functional basis of the occupational hierarchy. In A. C. Kerckhoff (Eds.), *Research in Sociology of Education and Socialization*, 5 (pp. 119-165). Greenwich, CT: JAI Press.
- Gottfredson, L. S. (1986). Societal consequences of the g factor in employment. *Journal of Vocational Behavior*, 29, 379-410.
- Gottfredson, L. S. (1988). Reconsidering fairness: A matter of social and ethical priorities. *Journal of Vocational Behavior*, 33, 293-319.
- Gottfredson, L. S. (1997). Why g matters: The complexity of everyday life. *Intelligence*, 24, 79-132.

- Gottfredson, L. S. (2002). Where and why *g* matters: Not a mystery. *Human Performance*, 15(1/2), 25-46.
- Granger, C. W. J. & Newbold, P. (1986). *Forecasting economic times series, 2nd ed.* San Francisco, CA: Academic Press.
- Grant, H., & Dweck, C. S. (2001). Cross-cultural responses to failure: Considering outcome attributions with different goals. In F. Salili, C. Y. Chui, & Y. Y. Hong (Eds.), *Student motivation: The culture and context of learning* (pp. 203–219). New York: Plenum Press.
- Griggs v. Duke Power Co. (1971). 401 U.S. 424, 91 Supreme Court 849, Certiorari to the United States Court of Appeals for the Fourth Circuit No. 124.
- Grudnik, J. L., & Kranzler, J. H. (2001). Meta-analysis of the relationship between intelligence and inspection time. *Intelligence*, 29, 523-535.
- Hagen, J. M., & Choe, S. (1998). Trust in Japanese interfirm relations: Institutional sanctions matter. *Academy of Management Review*, 23(3), 589-600.
- Hall, D. T. (1976). *Careers in organizations*. Pacific Palisades, CA: Goodyear.
- Hallilan, M. T., & Williams, R. A. (1990). Student characteristics and the peer-influence process. *Sociology of Education*, 63(2), 122-132.
- Hallman, S., & Ireson, J. (2005). Secondary school teachers' pedagogic practices when teaching mixed and structured ability classes. *Research Papers in Education*, 20, 3-24.
- Hanushek, E. A., Kain, J. F., Markman, J. M., & Rivkin, S. G. (2003). Does peer ability affect student achievement? *Journal of Applied Econometrics*, 18, 527–544.
- Herrnstein, R. J., & Murray, C. (1994). *The bell curve: Intelligence and class structure in American life*. New York: The Free Press.
- Herrnstein, R. J., Nickerson, R. S., Sanchez, M., & Swets, J. A. (1986). Teaching thinking skills. *American Psychologist*, 41, 1279-1289.
- Heugens, P. P. M. A. R., & Lander, M. W. (2009). Structure! Agency! (and other quarrels): A meta-analysis of institutional theories of organization. *Academy of Management Journal*, 52(1), 61-85.
- Huff, L. & Kelley, L. (2003). Levels of organizational trust in individualist versus collectivist societies: A seven-nation study. *Organization Science*, 14(1), 81-90.
- Hunter, J. E. (1980). *Validity generalization for 12,000 jobs: An application of synthetic validity and validity generalization to the General Aptitude Test Battery (GATB)*. Washington, D.C.: U.S. Employment Service, U.S. Department of Labor.
- Hunter, J. E. (1986). Cognitive ability, cognitive aptitudes, job knowledge, and job performance. *Journal of Vocational Behavior*, 29, 340–362.
- Hunter, J. E., & Schmidt, F. L. (1996). Intelligence and job performance: Economic and social implications. *Psychology, Public Policy, and Law*, 2, 447-472.
- Hunter, J. E., & Schmidt, F. L. (2004). *Methods of meta-analysis: Correcting error and bias in research findings* (2nd ed.). Thousand Oaks, CA: Sage.
- IBM. (2008). History of IBM: 1940s. Retrieved from the IBM Archives, http://www-03.ibm.com/ibm/history/history/decade_1940.html, December 28, 2008.
- Jensen, A. R. (1998). *The G factor: The science of mental ability*. Westport, CT: Praeger.
- Johns, G. (2006). The essential impact of context on organizational behavior. *Academy of Management Review*, 31, 386-408.
- Judge, T. A., Higgins, C. A., Thoresen, C. J., & Barrick, M. R. (1999). The big five personality traits, general mental ability, and career success across the life span. *Personnel Psychology*, 52, 621-652.

- Kambourov, G., & Manovskii, I. (2009). Occupational specificity of human capital. *International Economic Review*, 50 (1), 63-115.
- Kamin, L. J. (1974). *The science and politics of IQ*. Potomac, MD: Lawrence Erlbaum.
- Kanter, R. M. (1977). *Men and women of the corporation*. New York: Basic Books.
- Kerckhoff, A. C. (1993). *Diverging pathways: Social structure and career deflections*. New York/London: Cambridge Univ. Press.
- Kerckhoff, A. C. (1995). Institutional arrangements and stratification processes in industrial societies. *Annual Review of Sociology*, 15, 323-47.
- Kerckhoff, A. C. (2001). Education and social stratification processes in comparative perspective. *Sociology of Education*, 74, 3-18.
- Kerckhoff, A., & Glennie, E. (1999). The Matthew effect in American education. *Research in the Sociology of Education and Socialization*, 12, 35-66.
- Kevles, D. J. (1968). Testing the Army's intelligence: Psychologists and the military in World War I. *Journal American History*, 55, 565-581.
- Kim, K. I., & Organ, D. W. (1982). Determinants of leader-subordinate exchange relationships. *Group and Organization Studies*, 7, 77-89.
- Kram, K. E. (1983). Phases of the mentor relationship. *Academy of Management Journal*, 26, 608-625.
- Kuncel, N. R., Hezlett, S. A., & Ones, D. S. (2004). Academic performance, career potential, creativity, and job performance: Can one construct predict them all? *Journal of Personality and Social Psychology*, 86(1), 148-161.
- Landrum, R. E., Jeglum, E. B., & Cashin, J. R. (1994). The decision-making processes of graduate admission committees in psychology. *Journal of Social Behavior and Personality*, 9, 239-248.
- Landy, F. J. (1997). Early influences on the development of industrial and organizational psychology. *Journal of Applied Psychology*, 82, 467-477.
- Larsen, H. H., London, M., Weinstein, M., & Raghuram, S. (1998). High-flyer management-development programs. *International Studies of Management and Organization*, 28(1), 64-90.
- Law, K. S., Wong, C. S., Huang G. H., & Li, X. (2008). The effects of emotional intelligence on job performance and life satisfaction for the research and development scientists in China. *Asia Pacific Journal of Management*, 25, 51-69.
- Lemann, N. (1999). *The big test: The secret history of the American meritocracy*. New York: Farrar, Straus and Giroux.
- Levine, E. L., Spector, P. E., Menon, S., Narayanan, L., & Cannon-Bowers, J. (1996). Validity generalization for cognitive, psychomotor, and perceptual tests for craft jobs in the utility industry. *Human Performance*, 9, 1-22.
- Liden, R. C., Sparrowe, R. T., & Wayne, S. J. (1997). Leader-member exchange theory: The past and potential for the future. *Research in Personnel and Human Resource Management*, 15, 47-119.
- Liden, R., Wayne, S. J., & Stilwell, D. (1993). A longitudinal study on the early development of leader-member exchanges. *Journal of Applied Psychology*, 78, 662-674.
- Lievens, F. (2007). Research on selection in an international context: Current status and future directions. In M. M. Harris (Ed.) *Handbook of research in international human resource management* (pp. 107-123). New York: Lawrence Erlbaum Press.
- Lillard, L. A., & Tan, H. W. (1992). Private sector training: Who gets it and what are its effects? *Research in Labor Economics*, 13, 1-62.

- Lippmann, W. (1922). The abuse of tests. *New Republic*, 32, 297-298.
- Mackenzie, D. (2004). The big, bad wolf and the rational market: Portfolio insurance, the 1987 crash and the performativity of economics. *Economics and Society*, 33(3), 303-334.
- MacKenzie, D., Muniesa, F., & Siu, L. (2007). Introduction. D. MacKenzie, F. Muniesa, L. Siu, eds. *Do economists make markets? On the performativity of economics*. Princeton, NJ: Princeton Press.
- Madaus, G. F., Haney, W., & Kreitzer, A. (2000). The role of testing in evaluations. D. L. Stufflebeam, G. F. Madaus, T. Kellaghan, (Eds.). *Evaluation models: Viewpoints on educational and human services evaluation*, (pp. 113-126). Boston, MA: Kluwer Academic Publishers.
- Mahoney, J. (2000). Path dependence in historical sociology. *Theory and Society*, 29, 507-548.
- Marvin, W. M. (May 11, 1919). "Intelligence Tests." *The New York Times*: p. 36. Retrieved November 11, 2009 from <http://query.nytimes.com/mem/archive-free/pdf?res=9D02E3DD163AE03ABC4952DFB3668382609EDE>.
- Mathison, S. (2004). A short history of education assessment and standards-based educational reform. In D. Gabbard, E.W. Ross, eds. *Defending public schools*. Westport, CT: Praeger.
- Maume, D. J. (1999). Glass ceilings and glass escalators: Occupational segregation and race and sex differences in managerial promotions. *Work and Occupations*, 26, 483-509.
- McDaniel, M. A. (2006). Estimating state IQ: Measurement challenges and preliminary correlates. *Intelligence*, 34, 607-619.
- McHenry, J. J., Hough, L. M., Toquam, J. L., Hanson, M. A., & Ashworth, S. (1990). Project A validity results: The relationship between predictor and criterion domains. *Personnel Psychology*, 43, 335-355.
- Melby, J. N., Conger, R. D., Fang, S-A., Wickrama, K. A. S., Conger, K. J. (2008). Adolescent family experiences and educational attainment during early adulthood. *Developmental Psychology*, 44, 1519-1536.
- Mensh, E., & Mensh, H. (1991). *The IQ mythology: Class, race, gender, and inequality*. Carbondale, IL: Southern Illinois University Press.
- Merton, R. K. (1968). The Matthew effect in science. *Science*, 159, 56-63.
- Merton, R. K. (1988). The Matthew effect in science, II: Cumulative advantage and the symbolism of intellectual property. *ISIS*, 79, 606-623.
- Miller, D. T. (1999). The norm of self interest. *American Psychologist*, 54, 1053-1060.
- Miiller, W., & Shavit, Y. (1998). The institutional embeddedness of the stratification process: A comparative study of qualifications and occupations in thirteen countries. Y. Shavit & W. Miiller, (Eds.), *From school to work: A comparative study of educational qualifications and occupational destinations* (pp: 1-8). Oxford, UK: Clarendon Press.
- Mondale, S. (Director). (2001). *School: The story of American public education [Documentary]*. KCET/Hollywood: Public Broadcasting Service.
- Murphy, K. R. (Ed.) (2003). *Validity generalization: A critical review*. Mahwah, NJ: Erlbaum
- Murphy, K. R. (2009). Validity, validation and values. *Academy of Management Annals*, 3, 421 – 461.
- Murray, C. (1998). *Income inequality and IQ*. Washington, DC: AEI Press.
- National Association for College Admission Counseling (NACAC), 2006. State of College Admission. <http://www.nacacnet.org/NR/rdonlyres/4B4C0DF4-BF0A-4B10-A3D631061305/0/06StateofCollegeAdmissionpdf.pdf>. Retrieved March 8, 2008.

- Neal, D. (1995). Industry-specific human capital: Evidence from displaced workers. *Journal of Labor Economics*, 13, 653-677.
- Neisser, U., Boodoo, G., Bouchard, T. J., Boykin, A. W., Brody, N., Ceci, S. J., Halpern, D. F., Loehlin, J. C., Perloff, R., Sternberg, R. J., & Urbina, S. (1996). Intelligence: Knowns and unknowns. *American Psychologist*, 51(2), 77-101.
- Nisbett, R. E. (2009). *Intelligence and how to get it: Why schools and cultures count*. New York: Norton.
- Oakes, J. (1987). *Tracking in secondary schools: A contextual perspective*. Santa Monica, CA: Rand Corp.
- Pager, D., Western, B., & Bonikowski, B. (2009). Discrimination in a low wage labor market: A field experiment. *American Sociological Review*, 74, 777-99.
- Pomerantz, E. M., & Saxon, J. L. (2001). Conceptions of ability as stable self-evaluative processes: A longitudinal examination. *Child Development*, 72, 152-173.
- Rauschenberger, J. M., & Schmidt, F. L. (1987). Measuring the economic impact of human resource programs. *Journal of Business and Psychology*, 2(1), 50-59.
- Ratner, R. K., & Miller, D. T. (2001). The norm of self interest and its effects on social action. *Journal of Personality and Social Psychology*, 81, 5-16.
- Rawls, J. (2001). *Justice as fairness: A restatement*. Cambridge, Massachusetts: Belknap Press.
- Ree, M. J., Carretta, T. R., & Teachout, M. S. (1995). Role of ability and prior job knowledge in complex training performance. *Journal of Applied Psychology*, 80, 721-730.
- Ree, M. J., & Earles, J. A. (1992). Intelligence is the best predictor of job performance. *Current Directions in Psychological Science*, 1, 86-89.
- Ree, M. J., Earles, J. A., & Teachout, M. S. (1994). Predicting job performance: Not much more than g. *Journal of Applied Psychology*, 79, 518-524.
- Reeb, M. (1976). Differential test validity for ethnic groups in the Israel Army and the effects of educational level. *Journal of Applied Psychology*, 61, 253-261.
- Rindermann, H. (2007a). The g-factor of international cognitive ability comparisons: The homogeneity of results in PISA, TIMSS, PIRLS and IQ-tests across nations. *European Journal of Personality*, 21, 667-706.
- Rindermann, H. (2007b). Die Bedeutung der mittleren Klassenfahigkeit fur das Unterrichtsgeschehen und die Entwicklung individueller Fahigkeiten [The relevance of class ability for teaching and development of individual competences]. *Unterrichtswissenschaft*, 35, 68-89.
- Rindermann, H. (2008). Relevance of education and intelligence at the national level for the economic welfare of people. *Intelligence*, 36(2), 127-142.
- Roberson, L., & Block, C. J. (2001). Racioethnicity and job performance: A review and critique of theoretical perspectives on the causes of group differences. In B.M. Staw & R. I. Sutton (Eds.), *Research in Organizational Behavior*, 23, 247-325.
- Rogers, E. M. (2003) *Diffusion of Innovations* (5th ed.), New York: Free Press.
- Rosenbaum, J. E. (1979). Tournament mobility: Career patterns in a corporation. *Administrative Science Quarterly*, 24(2), 220-241.
- Rosenbaum, J. E. (1984). *Career mobility in a corporate hierarchy*, Orlando, FL: Academic Press.
- Roth, P. L., Huffcutt, A. I., & Bobko, P. (2003). Ethnic group differences in measures of job performance: A new meta-analysis. *Journal of Applied Psychology*, 88, 694-706.
- Rothstein, J. M. (2004). College performance predictions and the SAT. *Journal of Econometrics*, 121, 297 - 317.

- Rousseau, D. M., & Fried, Y. (2001). Location, location, location: Contextualizing organizational research. *Journal of Organizational Behavior*, 22, 1-13.
- Rowe, D. C., Vesterdal, W. J., & Rodgers, J. L. (1998). Herrnstein's syllogism: Genetic and shared environmental influences on IQ, education, and income. *Intelligence*, 26, 405-423.
- Ryan, A. M., McFarland, L., Baron, H., & Page, R. (1999). An international look at selection practices: Nation and culture as explanations for variability in practice. *Personnel Psychology*, 52, 359-391.
- Salgado, J. F., Anderson, N., Moscoso, S., Bertua, C., de Fruyt, F. (2003a). International validity generalization of GMA and cognitive abilities: A European community meta-analysis. *Personnel Psychology*, 56, 573-605.
- Salgado, J. F., Anderson, N., Moscoso, S., Bertua, C., de Fruyt, F., & Rolland, J. P., (2003b). A meta-analytic study of general mental ability validity for different occupations in the European community. *Journal of Applied Psychology*, 88, 1068-1081.
- Schmidt, F. (2002). The role of general cognitive ability and job performance: Why there cannot be a debate. *Human Performance*, 15(2), 187-202.
- Schmidt, F. L., & Hunter, J. E. (1977). Development of a general solution to the problem of validity generalization. *Journal of Applied Psychology*, 62, 643-661.
- Schmidt, F. L., & Hunter, J. E. (1986). Impact of job experience and ability on job knowledge, work sample performance, and supervisory ratings of job performance. *Journal of Applied Psychology*, 71(3), 432-439.
- Schmidt, F. L., & Hunter, J. E. (1992). Development of a causal model of processes determining job performance. *Current Directions in Psychological Science*, 1, 89-92.
- Schmidt, F. L., & Hunter, J. E. (1998). The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. *Psychological Bulletin*, 124, 262-274.
- Schmidt, F. L., & Hunter, J. E. (2000). Select on intelligence. E.A. Locke (Ed.), *Handbook of principles of organizational behavior*. Oxford, UK: Blackwell, 3-14.
- Schmidt, F. L., & Hunter, J. (2004). General mental ability in the world of work: Occupational attainment and job performance. *Journal of Personality and Social Psychology*, 86, 162-173.
- Schmidt, F. L., Hunter, J. E., McKenzie, R. C., & Muldrow, T. W. (1979). Impact of valid selection procedures on work-force productivity. *Journal of Applied Psychology*, 64, 609-626.
- Schmitt, N. (2007). The value of personnel selection: Reflections on some remarkable claims. *Academy of Management Perspectives*, 21, 19-23.
- Scullin, M. H., Peters, E., Williams, W. M., Ceci, S. (2000). The role of IQ and education in predicting later labor market outcomes: Implications for affirmative action. *Psychology, Public Policy and Law*, 6(1), 63-89.
- Secret mind tests of the Army; Ban lifted at last from censorship on mental exercises by which 1,500,000 soldiers were graded by experts on psychology for jobs that suited them. (1919, February 16). *New York Times*. Retrieved November 11, 2009 from <http://query.nytimes.com/mem/archivefree/pdf?res=9806E4DB1139E13ABC4E52DFB4668382609EDE> .
- Sewell, W. H., Haller, A. O., & Ohlendorf, G. W. (1970). The educational and early occupational status attainment process: Replication and revision. *American Sociological Review*, 35, 1014-1027.

- Shavit, Y., & Blossfeld, H-P. (Eds.). (1993). *Persistent inequality: Changing educational attainment in thirteen countries*. Boulder, CO: Westview.
- Shaywitz, B. A., Holford, T. R., Holahan, J. M., Fletcher, J. M., Stuebing, K. K., Francis, D. J., & Shaywitz, S. E. (1995). A Matthew effect for IQ but not for reading – Results from a longitudinal-study. *Reading Research Quarterly*, 30, 894-906.
- Sheridan, J. E., Slocum, Jr., J. W., Buda, R., & Thompson, R. C. (1990). Effects of corporate sponsorship and departmental power on career tournaments. *Academy of Management Journal*, 33, 578-602.
- Siegler, R. S. (1992). The other Alfred Binet. *Developmental Psychology*, 28(2), 179 – 190.
- Slan-Jerusalim, R., & Hausdorf, P. A. (2007). Managers' justice perceptions of high potential identification practices. *Journal of Management Development*, 26(10), 933 – 950.
- Stanovich, K. E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly*, 21(4), 360-407.
- Stanovich, K. E. (2000). *Progress in understanding reading: Scientific foundations and new frontiers*. New York: Guilford Press.
- Sternberg, R. J. (1988). *The triarchic theory of human intelligence*. New York: Viking-Penguin.
- Sternberg, R. J. (1998). Abilities are forms of developing expertise. *Educational Researcher*, 27, 11-20.
- Sternberg, R. J., Grigorenko, E. L., & Bundy, D. A. (2001). The predictive value of IQ. *Merill-Palmer Quarterly*, 47, 1-41.
- Stewart, E. B. (2008). School structural characteristics, student effort, peer associations, and parental involvement - The influence of school- and individual-level factors on academic achievement. *Education and Urban Society*, 40(2), 179-204.
- Terman, L. M (1916). *The measurement of intelligence: An explanation of and a complete guide for the use of the Stanford revision and extension of the Binet-Simon intelligence scale*. Boston, MA: Houghton Mifflin.
- Tests to ascertain fitness of workers. (1922, December 26). *The New York Times*. Retrieved November 11, 2009 from <http://query.nytimes.com/mem/archive-free/pdf?res=950CE0D71630E433A25755C2A9649D946395D6CF>
- Tsay, A., Lamont, M., Abbott, A., & Guetzkow, J. (2003). From character to intellect: changing conceptions of merit in the social sciences and humanities, 1951–1971. *Poetics*, 31, 23-49.
- Tsui, A. S., Nifadkar, S. S., & Ou, A.Y. (2007). Cross-national, cross-cultural organizational behavior research: Advances, gaps, and recommendations. *Journal of Management*, 33, 426-478.
- Turner, R. H. (1960). Sponsored and contest mobility and the school system. *American Sociological Review*, 25, 855-867.
- Walberg, H. J., & Tsai, S. (1983). Matthew effects in education. *American Educational Research Journal*, 20, 359-373.
- Weber, M. (1946). Bureaucracy. In H. Gerth & C.W. Mills (Eds.), *From Max Weber* (pp: 196-244). New York: Oxford University Press.
- Wernerfelt, B. (1984). A resource-based view of the firm. *Strategic Management Journal*, 5, 171-180.
- Wilk, S. L., & Cappelli, P. (2003). Understanding the determinants of employer use of selection methods. *Personnel Psychology*, 56(1), 103-124.
- Wilk, S. L., Desmarais, L. B., & Sackett, P. R. (1995). Gravitation to jobs commensurate with ability: Longitudinal and cross-sectional tests. *Journal of Applied Psychology*, 80, 79-85.

- Wilk S. L., & Sackett, P. R. (1996). Longitudinal analysis of ability-job complexity fit and job change. *Personnel Psychology*, 49, 937–967.
- Williams, W. M. (1998). *Are we raising smarter children today? School- and home-related influences on IQ and related measures*. Washington, DC: American Psychological Association.
- Willingham, W. W., & Breland, H. M. (1982). *Personal qualities and college admissions*. New York: College Entrance Examination Board.
- Yamagishi, T. (1988). The provision of a sanctioning system in the United States and Japan. *Social Psychology Quarterly*, 51, 265-271.
- Yamagishi, T., Cook, K. S., and Watabe, M. (1998). Uncertainty, trust and commitment formation in the United States and Japan. *American Journal of Sociology*, 104(1), 164–194.
- Yamagishi, T., & Yamagishi, M. (1994). Trust and commitment in the United States and Japan. *Motivation and Emotion*, 18, 129-166.
- Yang, R. (2004). Chapter 8. Toward massification: Higher education development in the People's Republic of China since 1949. In J.C. Smart (Ed.), *Higher education: Handbook of theory and research*: Vol. 19, (pp: 311–374). Netherlands: Springer.
- Yang, S-Y., & Sternberg, R. J. (1997). Taiwanese Chinese people's conceptions of intelligence. *Intelligence*, 25, 21–36.
- Young, M. (1958). *The rise of the meritocracy*. Edison, NJ: Transaction Publishers.
- Zagorsky, J. L. (2007). Do you have to be smart to be rich? The impact of IQ on wealth, income and financial distress. *Intelligence*, 35, 489-501.