Lang, J. W. B, Kersting, M. & Hülsheger, U. R.. (2010). Range shrinkage of cognitive ability test scores in applicant pools for german governmental jobs: Implications for range restriction corrections. International Journal of Selection and Assessment., 18, 321-328.

Range Shrinkage of Cognitive Ability Test Scores in Applicant Pools for German Governmental Jobs: Implications for range restriction corrections

Jonas W. B. Lang*, Martin Kersting** and Ute R. Hülsheger*

*Department of Work and Social Psychology, Maastricht University, Postbus 616, 6200 MD Maastricht, The Netherlands. jonas.lang@maastrichtuniversity.nl
**Federal University of Applied Administrative Sciences, Münster, Germany

Range restriction corrections require the predictor standard deviation in the applicant pool of interest. Unfortunately, this information is frequently not available in applied contexts. The common strategy in this type of situations is to use national-norm standard deviation estimates. This study used data from 8,276 applicants applying to nine jobs in German governmental organizations to compare applicant pool standard deviations for two cognitive ability tests with national-norm standard deviation estimates, and standard deviations for the total group of governmental applicants. Results revealed that job- and organizational context-specific applicant pool standard deviations were on average about 10–12% smaller than estimates from national norms, and about 4–6% smaller than standard deviations for the total group of governmental applicants.

1. Introduction

Professional guidelines (SIOP, 2003) and meta-analytic procedures (Hunter, Schmidt, & Le, 2006; Raju, Burke, Normand, & Langlois, 1991) advise applied researchers to correct job performance validity coefficients for statistical artifacts when the required information is available. Corrections for artifacts improve the precision of validity estimates so that validity coefficients approximate the true validity of a selection procedure more closely.

One important statistical artifact in the context of validation research is range restriction. Range restriction occurs when scores on a selection test are used to select persons from an applicant pool. As a result of the selection decision, the range of scores in the selected group of persons is smaller than the range of test scores in the original applicant pool. In validation research, only the information for the restricted sample of persons is available because only the selected persons actually get the opportunity to work in the job. Accordingly, validity coefficients can only be calculated for the restricted sample. However, in the context of validation research, researchers are interested in the validity of a selection procedure for the entire (unrestricted) applicant pool (the population of interest in validation research). In order to obtain an estimate of the predictive validity for the entire applicant pool, researchers need to use correction formulas (e.g., Hunter *et al.*, 2006; Raju *et al.*, 1991). These formulas require information on the predictor standard deviation in the restricted sample and the predictor standard deviation in the applicant pool of interest. Using this information, it is then possible to correct for selection effects and estimate the validity coefficient for the unrestricted applicant pool.

A major issue for researchers seeking to correct validity coefficients for range restriction is that the standard deviation in the applicant pool of interest is often difficult to obtain in applied contexts. Consequently, applied researchers frequently face situations, in which only information on the predictor standard deviation in the restricted sample is available (Kuncel & Klieger, 2007; Ones & Viswesvaran, 2003; Sackett & Ostgaard, 1994). The standard practice under these circumstances is to substitute missing applicant pool standard deviations with standard deviation estimates from national normative data. This practice is appropriate under the assumption that applicant pools are nonsystematic random samples from the general population. However, this assumption is not always tenable. Researchers have long been aware that standard deviations in applicant samples can potentially be smaller than standard deviation estimates from normative data because of self-selection effects (Jensen, 1980; Hartigan & Wigdor, 1989; Kuncel & Klieger, 2007; Ones & Viswesvaran, 2003; Sackett & Ostgaard, 1994). Self-selection effects occur when a significant proportion of potential applicants decide not to apply for a job and these persons have either higher or lower scores on the predictor than the persons who apply. High-ability persons may decide not to apply because they have other better choices (e.g., more attractive jobs). Persons with low-abilities, in contrast, may decide not to apply because they think that the probability of being hired is so low that applying is not worth the effort. Both high-ability and low-ability persons may also elect not to apply because they have a personal feeling that their person-job fit for the job in general or the recruiting organization is low (i.e., the job is too difficult or too easy). When self-selection effects are present, the variability in applicant pools is lower than the variability indicated in national normative data. Consequently, using normative data to substitute missing applicant pool standard deviations leads to an overestimation of the applicant pool standard deviation and, in consequence, to an overcorrection of the observed validity.

To determine the degree to which self-selection effects actually lower predictor standard deviations in applicant pools, researchers have begun to systematically study standard deviation differences between applicant pools and national norms (Kuncel & Klieger, 2007; Ones & Viswesvaran, 2003; Sackett & Ostgaard, 1994). The present study contributes to this emerging literature. We focus on range shrinkage on cognitive ability measures and extend previous research in two ways.

First, our study accounts for self-selection effects elicited by organizational characteristics. Previous research on applicant pool range shrinkage on cognitive ability measures has focused on job-specific self-selection effects by comparing job-specific applicant pools coming from a variety of different organizations and national norms (Sackett & Ostgaard, 1994). However, person–environment fit theories (e.g., Schneider, 1987) suggest that individuals with certain abilities might not only be attracted to specific jobs but also to specific organizations. Self-selection effects might therefore not only be job-specific, but also organization-specific. Our study accounts for this possibility by comparing job-specific applicant pools from a homogenous organizational context with national norms.

Second, we contribute to the literature by reporting data from the context of another nation. The previous research examined data for the North American version of an ability test (Sackett & Ostgaard, 1994). In contrast, we report data from Germany—a European nation with a different educational and vocational system and thus provide a critical test of the idea that range shrinkage in applicant pool standard deviations is a phenomenon that is not restricted to a specific instrument or a specific educational and vocational system.

2. Research on applicant pool range shrinkage

The first study examining restrictions of applicant pool standard deviations was conducted by Sackett and Ostgaard (1994). This study focused on applicant pool standard deviations of cognitive ability tests in the context of personnel selection. Sackett and Ostgaard examined data from the manual of the Wonderlic Personnel Test—a brief measure of general mental ability. Sackett and Ostgaard used the 1983 edition of the manual (E. F. Wonderlic & Associates Inc., 1983). This manual contains means and standard deviations of job-specific applicant pools for 80 jobs, which were tested with the US version of the Wonderlic Personnel Test. Sackett and Ostgaard compared this data with the standard deviation for the general workforce population. Because the 80 jobs covered a wide range of jobs and came from a diverse group of organizations, the general workforce standard deviation could be estimated by combining the standard deviations for the 80 job-specific applicant pools. Results revealed that applicant pool standard deviations were on average 8.3% smaller than the general workforce population estimate. Based on these findings, Sackett and Ostgaard recommended that researchers should not simply use data from national norms to substitute missing applicant pool standard deviations. Sackett and Ostgaard instead advised applied researchers to adjust standard deviations from national norms by a correction factor before correcting validity coefficient for range restriction.

After Sackett and Ostgaard conducted their initial study, a large body of new normative data and new job-specific applicant pool data for the Wonderlic Personnel Test was published (Wonderlic Personnel Test Inc., 1992 p. 27). We consequently examined whether the previous findings were still valid for this new data from 72 job-specific applicant pools. The findings were similar to those of the initial investigation with a slight tendency for larger differences between applicant pools and national norms. Applicant pool standard deviations were on average 11.57% smaller than the standard deviation for the general workforce population (a combination of the 72 job-specific applicant pools like in the initial study).

A second study by Ones and Viswesvaran (2003) extended Sackett and Ostgaard's work by examining applicant pool standard deviation for seven personality scales from an occupational personality inventory. In this investigation, applicant pool standard deviations were on averaged about 4% smaller than standard deviations estimates for normative data (range 2–9%). Ones and

Viswesvaran concluded that the use of estimates from normative data only slightly alters range restriction corrections for personality-criterion relationships in personnel selection and noted that their work generally supports Sackett and Ostgaard's approach of correcting standard deviation estimates for national normative samples downwards.

The most recent study comparing applicant group standard deviations and standard deviations of normative data focused on a boundary condition for self-selection effects (Kuncel & Klieger, 2007). The authors of this study focused on educational assessment. They studied standard deviations of an educational selection test—the Law School Admission Test. Specifically, the authors compared standard deviations in applicant pools for specific law schools with the standard deviation of all persons taking the test. The special circumstances of the study were that applicants knew their test scores and were able to retrieve information regarding law school rankings before they applied. Under these circumstances, Kuncel and Klieger (2007) found that applicant pool standard deviations were 23% smaller than the variability of the overall sample of test takers. Kuncel and Klieger's study illustrates that self-selection effects vary according to the circumstances of the respective study and can be quite substantial when applicants are provided with detailed information.

3. The present study

The present study focuses on the use of cognitive ability measures in the context of personnel selection. Our study seeks to extend previous research by accounting for self-selection effects elicited by organizational characteristics. The data examined in previous research in this area consisted of job-specific applicant pool standard deviations that combined all data for a specific job across a variety of organizational contexts. Consequently, this research provides information to what extent job-specific self-selection effects lead to shrinkage in standard deviations in applicant pools making them narrower than normative standard deviation estimates for the general workforce population.

Person–environment fit theories (Breaugh, 1992; Edwards, 2008; Schneider, 1987; Schneider, Goldstein, & Smith, 1995; Wanous, 1980) suggest that in addition to job-specific range restriction there might also be a second type of range restriction, namely organizational-context specific range restriction. Specifically, several person–environment fit theories argue that individuals tend to be attracted by organizations commensurate to their individual characteristics. This notion has most frequently been advanced for personality and interest constructs (e.g., Schneider, 1987; Schneider *et al.*, 1995), but several person–environment theories have extended this idea to the fit between abilities and demands (e.g., Breaugh, 1992; Wanous, 1980; Werbel & Gilliland, 1999). Accordingly, applicants may not only be interested in a specific type of job but may also consider organizational characteristics in their decision to apply for open positions. This notion has implications for the substitution of applicant pool standard deviations in range restriction corrections.

First, it is possible that organizational characteristics further contribute to shrinking applicant pool standard deviations for cognitive ability measures over and above effects observed in previous research comparing job-specific (but organizationally heterogenous) samples with national-norm standard deviation estimates. To investigate this possibility, we compared job-specific applicant standard deviations from a homogenous organizational context with national-norm standard deviation estimates.

Second, it is possible that previously documented range shrinkage in applicant pool standard deviations that has largely been interpreted as being job-related may in part result from organizational characteristics. To address this research question, we compared applicant standard deviations with the standard deviations for the total group of governmental applicants to determine the degree to which job characteristics contributed to range shrinkage among applicants that applied for jobs in the same organizational context. This investigation is also informative as many tests are only used within in a specific organizational context. Accordingly, normative data for these tests may also come from the specific organizational context and it is therefore relevant to investigate to what degree normative data of this type is affected by range shrinkage.

A noteworthy additional characteristic of our study is that it was conducted in the context of a different educational and vocational system. Germany is Europe's leading and the world's fourth largest economy. Personnel selection in Germany differs from countries that use North American tests in a variety of ways, which may influence applicant group variability of cognitive ability measures. The most notable differences include a different structure of the school system (Hülsheger, Maier, & Stumpp, 2007; Trautwein, Lüdtke, Marsh, Köller, & Baumert, 2006), a different structure of the vocational system (Muller, 1999; Schmidt & Foster, 1999), less ethnic and racial diversity (Salgado & Anderson, 2002; Viswesvaran & Ones, 2002), and a more similar view of personnel selection practices among personnel selection practitioners (Salgado *et al.*, 2003).

4. Method

4.1. Applicant pools

This study examined data from a total of 8,276 applicants in nine job-specific applicant pools. The data were collected by a large German testing agency that primarily works for the government of German states, districts, and cities. The applicants applied for on-the job training programs in one of nine different office or administration jobs. Germany has a special vocational training system which relies on formalized on-the job training programs for the majority of occupations (67% of the German workforce have been trained with on-the-job training programs of this type; Reinberg & Hummel, 2007). In this system, employees typically spend 70–80% of their time working on the job, where they are supervised by experienced supervisors with special training. The other 20–30% of the working time is reserved for formal training in governmental vocational schools (Muller, 1999; Schmidt & Foster, 1999). These type of on-the job training programs are accredited by the government and lasts 2–3 years. During this time, employees receive a lower salary and the term apprentice is attached to their job title. After they have finished their on-the-job training, employees get a certificate and a formal degree in the occupation.

4.2. *Cognitive ability tests*

We examined data for two cognitive ability tests. Both tests are part of the WIT-2 battery, which is a frequently used German ability-test battery (Kersting, Althoff, & Jäger, 2008). The first test used in the present study was a verbal analogy test consisting of 20 verbal analogies. Analogy tests of this type are frequently used in personnel selection and have been considered to be particularly effective as cognitive ability measures in personnel selection and educational assessment because they are relatively pure measures of general mental ability (Kuncel, Hezlett, & Ones, 2004). The second test used in the study was a vocabulary test containing 20 multiple-choice items. In this test, applicants read a word and must select a word that most closely matches it. Like analogy tests, vocabulary tests are saturated with general mental ability. In contrast, to analogy tests, however, they do not demand considerable amounts of information processing or reasoning. Instead, most authors argue that vocabulary tests primarily measure acquired knowledge or crystallized intelligence (Beier & Ackerman, 2005).

4.3. *National-norm standard deviation estimates*

To compare standard deviations and means from job-specific applicant pools with general workforce norms, we used the norm estimates reported in the test manual. Normative mean and standard deviation estimates for the two tests are $M = 7.88$ and $SD = 3.98$ for the verbal analogy test, and $M = 9.85$ and $SD = 4.00$ for the vocabulary test, respectively. These estimates are based on normative information from German workforce data, which is weighted in accordance with the education distribution of the German workforce population. The workforce samples used to calculate the education-weighted normative information consist of 9,219 and 10,024 persons and incorporate the data that we used for this study (Kersting *et al.*, 2008).

5. Results

Tables 1 and 2 provide details for the nine applicant pools in the present study. Table 1 provides the job title for each of the nine jobs. To document the exact nature of the nine jobs, we also provide the corresponding job codes in the BERUFENET classification of the German Working Agency (Bundesagentur für Arbeit, 2009) and corresponding job codes for the American Occupational Information Network (O*NET) job classification system (O*NET Online, 2009). Table 2 then provides applicant-pool specific means, standard deviations for the two cognitive ability tests.

5.1. *Variability of applicant means and standard deviations*

Before we examined range shrinkage, we first examined basic evidence for the presence of self-selection effects elicited by job characteristics. Specifically, we investigated whether means and standard deviations varied in the nine applicant pools. If more capable employees strive for certain jobs, while less capable employees tend to apply for other jobs, one should expect that some applicant pools are more homogenous than others and that the mean-ability level of the applicants is higher for some jobs than for others.

To examine the degree to which means were similar across the applicant pools, we examined the ICC1 (Bliese, 2000). The ICC1 indicates how much of the variability in test scores can be explained by group membership (i.e., applicant pool membership) and can be estimated from a one-way random-effects analysis of variance (ANOVA) model (see Bliese, 2000, for details).

Table 1. Job titles and corresponding job descriptions in BERUFENET and O*NET

Job title	BERUFENET code	O*NET
Municipal clerks	B 7811-902, B 7811-921	43-4031.02
Administrative supervisors	B 7621-928	43-1011.00
Administrative assistants	B 7811-952	43-9022.00
Correspondence clerks	B 7811-904	43-4021.00
File clerks	B 8234-904, B 8234-905, B 8234-906	43-4071.00
Computer software engineers and programmers	B 7748-917	15-1021.00, 15-1031.00, 15-1032.00
Office managers	B 7810-901	43-6011.00
Computer and information systems managers	B 7746-911	11-3021.00
Insurance clerks	B 7811-908	43-9041.00

Note. O*NET = Occupational Information Network (O*NET) SOC 2006 Code.

Table 2. Applicant pool means, applicant pool standard deviations, ratios of specific applicant pool standard deviations to normative standard deviation estimates (norm SD_{ratios}), and ratios of specific applicant pool standard deviations to standard deviations of the total group of governmental applicants (total group SD_{ratios})

Job title	*n*	Verbal analogy test					Vocabulary test					r_{M1-M2}
		α	*M*	*SD*	SD_{ratio}		α	*M*	*SD*	SD_{ratio}		
					Norm	Total group				Norm	Total group	
Municipal clerks	3,723	.74	7.33	3.66	0.92	0.98	.75	8.47	3.61	0.90	0.97	.55
Administrative supervisors	31	.78	7.97	3.47	0.87	0.93	.73	12.03	3.48	0.87	0.93	.40
Administrative assistants	2,190	.78	7.61	3.91	0.98	1.05	.77	9.36	3.87	0.97	1.04	.58
Correspondence clerks	113	.70	6.90	3.31	0.83	0.89	.70	7.98	3.31	0.83	0.89	.59
File clerks	162	.77	8.87	3.87	0.97	1.04	.77	10.40	3.89	0.97	1.04	.62
Computer software engineers and programmers	601	.74	9.23	3.73	0.94	1.00	.71	9.94	3.46	0.87	0.93	.57
Office managers	128	.70	6.80	3.36	0.84	0.90	.72	7.88	3.37	0.84	0.90	.58
Computer and information systems managers	36	.68	8.97	3.36	0.84	0.90	.57	10.31	2.95	0.74	0.79	.36
Insurance clerks	1,292	.77	6.44	3.66	0.92	0.98	.73	9.28	3.61	0.90	0.97	.40

Note. α = Cronbach's α coefficient. Norm $SD_{ratio} = (SD_{specific\ applicant\ pool}/SD_{manual\ norms})$; Total group $SD_{ratio} = (SD_{specific\ applicant\ pool}/SD_{total\ group})$. r_{M1-M2} = correlation between the two tests. Total group standard deviations were calculated by using the analysis of variance (ANOVA) theorem (i.e., total variance equals the variance of the means plus the mean of the variances). Because of the large sample-size of some applicant pools, we did not use sample-size based weighting so that each applicant pool received a similar weight.

We found ICC1 = .027; $F(8, 8{,}267) = 26.88$, $p < .001$, for the verbal analogy test, and ICC1 = .035; $F(8, 8{,}267) = 34.33$, $p < .001$, for the vocabulary test. These values are somewhat smaller than ICC1 values commonly reported in the literature (on average about .12, see James, 1982, and commonly ranging from .05 to .20, see Bliese, 2000). Values reported in the literature, however, commonly come from groups that already work together for a considerable time. In contrast, the similarity between the applicants in the present data set entirely steams from self-selection and the data consequently suggests that self-selection effects contribute to a considerable amount of similarity in the group means.

To examine the degree to which standard deviations were similar across the applicant pools, we conducted Bartlett tests of variance homogeneity. These tests also revealed that there were significant differences in variances between applicant pools (verbal analogy test: Bartlett's $K^2 = 21.11$, $p < .01$; vocabulary test: Bartlett's $K^2 = 27.89$, $p < .001$). It is relevant to note, however, that (like for the ANOVA-based analyses) the small *p*-values do not necessarily suggest that the effects are large because the sample size is relatively large. The analyses nevertheless provide some evidence that the variance of the applicant pools was reduced because of self-selection.

5.2. *Applicant pool standard deviations and normative (general workforce) standard deviations*

To investigate how strongly job-specific and organizational-context specific applicant pool standard deviations differed from normative standard deviation estimates, we computed standard deviation ratios (SD_{ratios}) by dividing applicant pool standard deviations by the normative standard deviation estimate for the respective test. Table 1 provides SD_{ratio} values for the nine applicant pools. Note that an SD_{ratio} of 1.00 results from equal standard deviations for job-specific applicant pools and national norms. Values smaller than 1.00 indicate that the applicant pool standard deviation is smaller, and values larger than 1.00 indicate that the applicant pool standard deviation is larger than the normative estimate.

For the verbal analogy test, we found an average SD_{ratio} of .90 (95% CI: .86, .95) indicating that applicant pool standard deviations were on average 10% smaller than the normative standard deviation (for SD_{ratios} for each applicant pool, see Table 2). The average SD_{ratio} for the vocabulary test was .88 (95% CI: .82, .93) so that applicant pool standard deviations were 12% smaller than the normative standard deviation. To examine whether these differences were significant, we conducted one sample *t*-tests testing the hypothesis that SD_{ratios} were < 1. Results revealed that the SD_{ratio} values for both tests were significantly < 1, verbal analogy test: $t(8) = -5.08$, $p < .001$; vocabulary test: $t(8) = -5.08$, $p < .001$.

5.3. *Applicant pool standard deviations and standard deviations for the total group of governmental employees*

We finally compared applicant pool standard deviations with standard deviations for the total group of governmental employees. Analyses were conducted by first computing standard deviations for the total group of governmental employees using the ANOVA theorem (i.e., total variance equals the variance of the means plus the mean of the variances). Because of the large sample-size of some applicant pools, we did not use sample-size based weighting so that each applicant pool received a similar weight. In the next step, we estimated SD_{ratios} by dividing applicant pool standard deviations by the total group standard deviations.

Analyses revealed that the total group standard deviation values for the two tests were about 6–7% smaller than the normative standard deviations, verbal analogy test: $SD = 3.73$ (national-norm estimate: $SD = 3.98$); vocabulary test: $SD = 3.73$ (national-norm estimate: $SD = 4.00$). People that apply for governmental jobs thus form a considerable more homogenous group than the general workforce norm group. SD_{ratios} describing the ratio of applicant pool standard deviations to total group standard deviations are provided in Table 2. For the verbal analogy test, we found an average SD_{ratio} of .96 (95% CI: .92, 1.01) indicating that applicant pool standard deviations were on average 4% smaller than the total group standard deviation value. This difference was not significant at the 5% level, $t(8) = -1.78$, $p = .06$. For the vocabulary test, the average SD_{ratio} was .94 (95% CI: .88, 1.00) suggesting that applicant pool standard deviations were on average 6% smaller than the total group standard deviation value for the vocabulary test. This difference was significant, $t(8) = -2.34$, $p = .03$.

6. Discussion

The present study replicates previous findings indicating that applicant pool standard deviations of cognitive ability measures are moderately smaller than normative standard deviation estimates. We found that applicant pool standard deviations were on average 10–12% smaller than estimates based on normative data. These values are close to the 8% value Sackett and Ostgaard (1994) found in their initial study with data from the 1983 Wonderlic Manual, and the 12% value for the data from the 1992 edition of the Wonderlic Manual. Our study shows that range shrinkage in applicant pool standard deviations is a phenomenon, which generalizes beyond the Wonderlic test, and consequently provides additional evidence that applicants actively consider their chances before they apply for jobs.

A novel contribution of our investigation is that we examined not only job-specific applicant pools but applicant pools which were also organizational context specific. This allowed us to differentiate between range restriction due to job-specific and due to organization-specific self-selection. The analyses suggested that range shrinkage was to a considerable degree organization-specific. About 4–6% of the range shrinkage we observed was clearly job-specific because it resulted from governmental job applicants' decision to apply for a specific type of job within a governmental institution. The remaining amount of range shrinkage generalized across applicants applying for governmental jobs and can thus be attributed to an organization-specific self-selection effect. These findings suggest that organizational characteristics may play an important role in addition to specific characteristics of jobs in people's decision to apply. This is in line with the person–environment fit literature suggesting that individuals choose to work in environments that match not only their personality characteristics and interests but also their abilities (e.g., Breaugh, 1992; Wanous, 1980; Werbel & Gilliland, 1999).

The present study has several implications. One implication of our study is that the findings provide applied researchers with a more solid basis for making decisions on the use and the adjustment of normative data when applicant pool data is missing for range restriction corrections. Sackett and Ostgaard (1994) noted that it may be reasonable to use research on range shrinkage in applicant pool standard deviations to correct normative data downward by a correction factor before one uses normative data for range restriction corrections. Thereby, researchers can prevent overcorrection of observed validities that lead to an overestimation of operational validities. Sackett and Ostgaard proposed that researchers may either use the average range shrinkage value or, alternatively, they may rely on a lower bound estimate like the 90th percentile in order to be conservative. Although the number of applicant pools in our data set was small and percentile values accordingly need to be treated with some caution, the 90% percentile values for the present data (16% for the verbal analogy test, and 19% for the vocabulary test) were nevertheless close to the value found by Sackett and Ostgaard (20%). In our German data, applicant pool range shrinkage thus occurred at a similar level as in the data gathered with the US version of the Wonderlic Personnel Test. Consequently, there is now a basis for assuming that range shrinkage in cognitive ability test scores is a phenomenon which is not restricted to data gathered with the US version of the Wonderlic Personnel Test. It rather seems to generalize to other cognitive ability measures and to other countries. Our study thereby adds to recent developments in personnel selection research to explicitly address and test cultural differences in personnel selection and expand the focus to European countries. After researchers have provided evidence on the generalizability of the validity of cognitive ability measures (e.g., Hülsheger *et al*., 2007; Salgado *et al*., 2003) and the generalizability of applicant reactions to selection procedures (Anderson & Witvliet, 2008), we now provide initial evidence for the generalizability of range restriction in applicant pools.

Additional implications of the present research relate to our finding that the total sample of governmental employees was considerably more homogenous than the normative data. This finding clearly has implications for tests that are used and normed in restricted organizational contexts instead of the general workforce. When researchers use normative data for instruments of this type, range shrinkage is likely a considerably smaller issue and the use of corrections is likely less frequently required. Likewise, when applicant data for other jobs in the same organizational context is available, researchers may consider using this context-specific data when applicant pool data is missing instead of normative information form test manuals. Using context-specific data may help preventing overcorrections of observed validities and may lead to more accurate estimates of the validity of selection measures.

In closing, we acknowledge limitations of our investigation. Specifically, it is unclear whether our finding of organizational-context specific range restriction is specific to governmental organizations. It is possible that only governmental organizations attract applicants with specific abilities, whereas other organizations might not elicit such self-selection effects. Future research may therefore address the generalizability of the organizational self-selection effect found in the present study.

In conclusion, we extended previous research by collecting new evidence on applicant pool shrinkage of cognitive ability scores in personnel selection. Specifically, this study replicates previous research showing that applicant pool standard deviations are smaller than normative standard deviations and show that this range shrinkage does not entirely result from job characteristics but also from organizational characteristics.

References

Anderson, N. and Witvliet, C.I. (2008). Fairness reactions to personnel selection techniques: An International comparison between the Netherlands, the United States, France, Spain, Portugal, and Singapore. *International Journal of Selection and Assessment*, *16*, 1–13.

Beier, M.E. and Ackerman, P.L. (2005). Age, ability, and the role of prior knowledge on the acquisition of new domain knowledge: Promising results in a real-world learning environment. *Psychology and Aging*, *20*, 341–355.

Bliese, P.D. (2000). Within-group agreement, non-independence, and reliability: Implications for data aggregation and analysis. In Klein, K.J. and Kozlowski, S.W. (Eds.), *Multilevel theory, research, and methods in organizations: Foundations, extensions and new directions* (pp. 349–381). San Francisco, CA: Jossey-Bass.

Breaugh, J.A. (1992). *Recruitment: Science and practice*. Boston, MA: PWS-Kent.

Bundesagentur für Arbeit. (2009). BERUFENET. Retrieved from http://berufenet.arbeitsagentur.de/

Edwards, J.R. (2008). Person–environment fit in organizations: An assessment of theoretical progress. *The Academy of Management Annals*, *2*, 167–230.

Hartigan, J.A. and Wigdor, A.K. (Eds.). (1989). *Fairness in employee testing: Validity generalization, minority issues, and the General Aptitude Test Battery*. Washington, DC: National Academy Press.

Hülsheger, U.R., Maier, G.W. and Stumpp, T. (2007). Validity of general mental ability for the prediction of job performance and training success in Germany: A meta-analysis. *International Journal of Selection and Assessment*, *15*, 3–18.

Hunter, J.E., Schmidt, F.L. and Le, H. (2006). Implications of direct and indirect range restriction for meta-analysis methods and findings. *Journal of Applied Psychology*, *91*, 594–612.

James, L.R. (1982). Aggregation bias in estimates of perceptual agreement. *Journal of Applied Psychology*, *67*, 219–229.

Jensen, A.R. (1980). *Bias in mental testing*. New York, NY: Free Press.

Kersting, M., Althoff, K. and Jäger, A.O. (2008). *Wilde-Intelligenz-Test 2 (WIT-2) [Wilde Intelligence Test 2 (WIT-2)]*. Göttingen, Germany: Hogrefe.

Kuncel, N.R., Hezlett, S.A. and 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*, 148–161.

Kuncel, N.R. and Klieger, D.M. (2007). Application patterns when applicants know the odds: Implications for selection research and practice. *Journal of Applied Psychology*, *92*, 586–593.

Muller, M. (1999). Human resource management under institutional constraints: The case of Germany. *British Journal of Management*, *10*, 31–44.

Ones, D.S. and Viswesvaran, C. (2003). Job-specific applicant pools and national norms for personality scales: Implication for range restriction corrections in validation research. *Journal of Applied Psychology*, *88*, 570–577.

Occupational Information Network (O* NET) Online. (2009). O*NET 10.0 database. Retrieved from http://online.onetcenter.org/

Raju, N.S., Burke, M.J., Normand, J. and Langlois, G.M. (1991). A new meta-analytic approach. *Journal of Applied Psychology*, *76*, 432–446.

Reinberg, A. and Hummel, M. (2007). *Qualifikationsspezifische Arbeitslosigkeit im Jahr 2005 und die Einführung der Hartz-IV-Reform. [Qualification-specific unemployment and the implementation of the Hartz-IV reform]. (IAB-Forschungsbericht 9/2007)*. Nürnberg, Germany: Bundesanstalt für Arbeit Retrieved from http://doku.iab.de/forschungsbericht/2007/fb0907.pdf

Sackett, P.R. and Ostgaard, D.J. (1994). Job-specific applicant pools and national norms for cognitive ability tests: Implications for range restriction corrections in validiation research. *Journal of Applied Psychology*, *79*, 680–684.

Salgado, J.F. and Anderson, N. (2002). Cognitive and GMA testing in the European community: Issues and evidence. *Human Performance*, *15*, 75–96.

Salgado, J.F., Anderson, N., Moscoso, S., Bertua, C., de Fruyt, F. and Rolland, J.P. (2003). A meta-analytic study of general mental ability validity for different occupations in the European community. *Journal of Applied Psychology*, *88*, 1068–1081.

Schmidt, K. and Foster, P.N. (1999). Germany's dual vocational training system. *Tech Directions*, *57*, 15–17.

Schneider, B. (1987). The people make the place. *Personnel Psychology*, *40*, 437–453.

Schneider, B., Goldstein, H.W. and Smith, D.B. (1995). The ASA framework: An update. *Personnel Psychology, 48*, 747–773.
Society for Industrial and Organizational Psychology (SIOP). (2003). *Principles for the validation and use of personnel selection procedures* (4th ed.). Bowling Green, OH: Author.
Trautwein, U., Lüdtke, O., Marsh, H.W., Köller, O. and Baumert, J. (2006). Tracking, grading, and student motivation: Using group composition and status to predict self-concept and interest in ninth-grade mathematics. *Journal of Educational Psychology, 98*, 788–806.
Viswesvaran, C. and Ones, D.S. (2002). Agreements and disagreements on the role of general mental ability (GMA) in industrial, work, and organizational psychology. *Human Performance*, *15*, 211–231.
Wanous, J.P. (1980). *Organizational entry: Recruitment, selection, and socialization of newcomers*. Reading, MA: Addison-Wesley.
Werbel, J.D. and Gilliland, S.W. (1999). Person–environment fit in the selection process. In Ferris, G.R. (Ed.), *Research in personnel and human resources management* (*Vol. 17*, pp. 209–243). Stamford, CT: JAI Press.
Wonderlic, E.F. and Associates Inc. (1983). *Wonderlic personnel test manual*. Northfield, IL: Author.
Wonderlic Personnel Test Inc. (1992). *Wonderlic personnel test user's manual. Manual*. Libertyville, IL: Author.