Thirteen years after. Using hierarchical linear modeling to investigate long-term assessment center validity

Paul Jansen (Vrije Universiteit Amsterdam, The Netherlands)

Thirteen years after. Using hierarchical linear modeling to investigate long-term assessment center validity

Abstract

Using hierarchical linear modeling the author investigated temporal trends in the predictive validity of an assessment center for career advancement (measured as salary growth) over a 13-year period, for a sample of 456 academic graduates. Using year of entry and tenure as controls, linear and quadratic properties of individual salary curves could be predicted by the assessment center dimensions. The validity of the (clinical) overall assessment rating for persons with tenure of at least 12 years corrected for initial differences in starting salaries, and restriction in range was .35. Corrected for predictor unreliability it became .40. In particular substantial long-term validities were obtained for the group discussion and analysis/presentation exercises. In accordance with findings from research in managerial effectiveness and development, the validity of the dimension firmness increased with tenure. Validity trends for two types of interviews and a mental test were also studied. The analytical and verbal ability dimensions of the mental test had negative long-term validities.

Thirteen years after. Using hierarchical linear modeling to investigate long-term assessment center validity

Jansen & Stoop (2001) investigated temporal trends in the validity of an assessment center consisting of a group discussion and an analysis/presentation exercise, for predicting career advancement as measured by average salary growth obtained over a 7-year period, for a sample of 679 academic graduates. The assessment center was part of a larger procedure in which also a mental test and two interviews (with a 'recruitment officer' and with a top-level manager) were incorporated. The validity of the overall assessment rating for persons with tenure of 7 years, corrected for initial differences in starting salaries and for restriction in range, was .39. There was a considerable time variation in the validity of both the overall assessment rating and assessment center dimensions. In accordance with findings from research in managerial effectiveness and development, the dimension interpersonal effectiveness only became valid after a number of years, while the dimension firmness was predictive in the whole period and increased in time. For comparison, validity trends were also investigated for the mental test and the two interviews.

In this study we present results of a continuation of this longitudinal study. Additional salary criterion data were collected up to 13 years after entry. Due to the split-of of a major part of the company in 1997, we had to limit the longitudinal study to the remaining 456 subjects. Data will be analyzed by hierarchical linear modeling in order to examine the existence, nature, and causes of within and between-person changes in salary over time (cf. Deadrick, Bennett & Russell, 1997).

DETERMINANTS OF OBJECTIVE CAREER SUCCESS

Long-term assessment center validity

Studies on the long-term validity of assessment centers for 'objective' advancement criteria such as career progress and salary advancement have produced mixed results. Validity decreased (Tziner, Ronen & Hacohen, 1993), increased (Anstey, 1977; Hinrichs, 1978; McEvoy & Beatty, 1989; Mitchel, 1975; Moses, 1971: see Huck, 1977), followed an inverted-U shaped trend (Thornton & Byham, 1982), or was independent of time (Gaugler et al., 1987). So it is unclear whether the

generally observed decline in predictive validity of selection instruments (Hulin, Henry & Noon, 1990; Keil & Cortina, 2001) also is the case for the assessment center (Barrett, Alexander & Doverspike, 1992). An explanation for varying long-term assessment center validities is, that career success factors change with tenure, and therefore the validities of assessment center dimensions predicting those factors also change over time (Hulin et al., 1990).

Both longitudinal studies into assessment center validity and research into (management) development (Fiedler & House, 1994; Hogan, Curphy & Hogan, 1994), suggest that the following dimensions (denotations in italic refer to dimensions used in the present study) are important for career progression:

- *'thinking':* general mental ability, analytical reasoning, creativity, openness;

- '*interpersonal effectiveness'* : networking, socially oriented and capable, interpersonally sensitive and competent, 'Extraversion from the 'Big Five';

- *'firmness'* (inverse of 'Neuroticism' from the 'Big Five'): autonomous, decisive, perseverance, able to cope with stress, dominance, self-confidence;

- '*ambition*': involvement, achievement motivation, desire to get ahead, commitment, initiative, energy level, drive,

and that '*operational competence'* (planning and organizing, systematic) becomes less important. For example, Ferris et al. (2001) found that general mental ability and social skills were predictive of salary level. In the study by Seibert, Crant & Krainer (1999), salary growth was associated with a 'proactive personality', a stable disposition consisting of autonomy, initiative, and perseverance. These comprise the 'social capital' that enhances objective career advancement. Ambition appears an important predictor of executives' objective career success in the sense of pay increases and promotion (Hogan & Holland, 2003; Judge et al., 1995). Leadership is related to energy, dominance, sociability, initiative, persistence, openness and creativity (Judge et al. 2002). However, Seibert & Krainer (2001) found, in a broad study with a diverse set of organizations and occupations, a negative relation between openness and salary level. Extraversion and firmness were related positively to salary level. Recent studies show operational competence (planning, conscientiousness) to be unrelated (Boudreau, Boswell & Judge, 2001) or related negatively (Robertson, 2000) to promotion into management. Generally, it seems that persuasive and

initiating behaviors ('firmness' and 'ambition') are a constant determinant of management progress, that the impact of interpersonal behaviors ('interpersonal effectiveness') gradually increases, and that operational behavior ('operational competence') can in time become a contra-indicator. The impact of general mental ability ('thinking'), and in particular openness, is less clear.

Predicting salaries curves from selection ratings

In our previous study, average salary growth: the difference between last salary and starting salary divided by total tenure, was taken as criterion. But this is only a limited indication of career advancement since we have per person not only starting salary and final salary, but also all salaries for the years in between. We therefore took in the present study as criterion properties of the individual salary *curve*. For each subject, his/her salary earned at 31 December was registered (in the company at issue all salary changes took place at 1 January, independent of the date of entry). Then, salary $y_{(ij)}$ of subject i in the j-st (calendar) year of the career can be written in curvilinear form as:

$$y_{(ij)} = \beta_{(0i)} + \beta_{(1i)}.(j-1) + \beta_{(2i)}.(j-1)^2 + r_{(ij)}, \tag{1}$$

where $\beta_{(0i)}$ is the constant, $\beta_{(1i)}$ the linear regression weight (slope), $\beta_{(2i)}$ the quadratic regression weight, and $r_{(ij)}$ the residual error due to nonlinear and non-quadratic salary trends. For instance, $y_{(i5)}$ is the salary person i earned in the 5th (calendar) year of the career. For someone whose career started at some moment in 1989, j ranges from 1 (salary earned at 31 December 1989) to 14 (salary earned at 31 December 2002). The initial or starting salary $y_{(i1)}$ is equal to $\beta_{(0i)}$. On account of a visual inspection of the curves only linear and quadratic salary trends were computed. Moreover, the literature referred to earlier supports a linear model in salary trend on account of the predictors.

In terms of hierarchical linear modeling ('HLM'; Hofmann, 1997), equation (1) specifies a within-person model of intra-individual variability of salary level over time. In the first, within-subjects part of the HLM-analysis ('level-1'), the parameters of the individual salary curves will be estimated in order to find out whether there are systematic, non-random inter-individual differences in salary changes over time. To find out whether individual differences as measured in the selection phase account for systematic salary

change patterns, the parameters of the individual salary curves will in the second, between-subjects part of the HLM-analysis ('level-2') be regressed on (1) year of entry and total tenure (controls), and (2) the selection scores (predictors).

We incorporated in the level-2 analysis year of entry as control because when the time period is that long, it is possible that cohort effects appear. That is, it could be that the group that for example started their career in 1989 differs, in selection profile or career conditions (for instance an accidental shortage of upper management positions) from the group that started in 1994, and that consequently the pattern of predictive validities changes. Cohort-effects can be caused by changes in the persons recruited, in the selection and promotion procedures, or in the organization. Comparing biographical, selection and promotion data for the cohorts can assess the first and second causes. In the analysis, we have to assume that the organization remains more or less unchanged in the period investigated. There was one major organizational change in the present study: in 1997 a large business unit became an independent company and since then developed and applied its own policies and procedures for personnel selection, appraisal and remuneration. While this resulted in an increase in missing data, it did not affect corresponding personnel policies or procedures for the remainder of the organization. Also, it is difficult to imagine a relationship between this separation and individual selection or criterion data (but we will check his). Finally, we also included total tenure as a control in the level-2 regression analysis. In this way, not absolute salary level but average salary growth was the dependent variable (just as in our first study)

We will also compare the long-term predictive validity of the assessment center to such other assessment instruments as the mental test and the selection interview. Since turnover is higher for low performers (Sturman & Trevor, 2001), we investigate the relation between both the predictors and the criterion on the one hand, and turnover on the other hand.

METHOD

Subjects

Data were collected from the privatized Netherlands Postal and Telecommunications Services 'KPN'. The study originally involved 679 academic graduates from Dutch universities recruited for a career

in management in the years 1989-1996 by KPN. Since the split-off of a major business unit in 1997

resulted in a loss of 223 subjects (33%), 456 persons were involved in the study. Criterion data were

collected from 1989-2002. Persons did not have any significant prior work experience. The academic

background of the recruits was (business) economics (about 35%), business administration and

management science (about 35%), engineering, technics and computer science (about 25%), and

miscellaneous (e.g. law, history, or social science). We compare predictor data collected during the

selection procedure with salary criterion data for 1-13 years after time of hire. Since last criterion data

were obtained at December 2002, we even had data for total tenure of 14 years (=subjects who started

somewhere in 1989), but the number of subjects with this tenure (5) turned out to be too low.

Description of selection procedure

Since the selection procedure was described at length in Jansen & Stoop (2001) we will limit

the description to a brief overview. The following general dimensions were assessed:

-*'thinking'*: intelligence, cognitive functioning in all its aspects: analytical reasoning, problem solving

capability, creativity, imagination;

-*'interpersonal effectiveness*: socially oriented and capable, interpersonally sensitive and competent,

being open towards others and being able to deal with others, 'extraversion' from the 'Big Five';

-*'firmness*': independent, strong, decisive, resistant, stamina, able to cope with stress, dominance, self-

confidence;

-'*ambition*': involvement, achievement motivation, commitment, energy level, drive.

-'*operational competence*': planning and organizing, productive, effective, systematic.

The selection procedure consisted of five consecutive and selective steps: a. Selection based on the

applicant's letter, by application of formal criteria as for instance field of study. b. Interview conducted

by a 'recruitment officer' (the 'recruitment interview'). c. Mental test. d. Interview with the manager

who is in charge of the department where the candidate will start his or her career ('management

interview'). e. Assessment center, consisting of a group discussion, an analysis/presentation exercise,

and an end meeting in which an overall assessment end rating (OAR) is established. Candidates can

be rejected at every step. Even when the end result of a step turns out to be positive, candidates can

withdraw from the procedure. After every step the candidate gets immediate feedback about his/her admission to the next step. But, knowledge obtained about a candidate at a previous step is never transmitted to the assessors who participate at a later step. In this study we concentrate on steps b-e.

Step b: Recruitment interview. 'Recruitment officers' of the company rate the candidate on the five dimensions, and also give an overall rating. Next they decide on the continuance of the selection procedure.

Step c: Mental test. The mental test consisted of nine paper and pencil tests. The tests result in one overall final score, and four so-called 'factor scores': on numerical ability, analytical ability (general reasoning), verbal ability, and creativity ('divergent production' in the sense of Guilford, 1967: velocity and productivity of verbal association). A cutoff score was determined for the final score. Persons with a final test score below the cutoff score were rejected.

Step d: Management interview. In the management interview the senior manager who is the 'owner' of the vacancy and consequently may become the manager of the graduate, decides on both the fit of the candidate for the job at issue, and his/her potential for management development. Assessment dimensions and rating procedures were the same as in the recruitment interview.

Step e: Assessment center. The final part of the procedure is an assessment center consisting of two situational exercises followed by an end meeting in which the OAR is determined. Up to this stage, candidates do not have contact with other candidates. However in the final assessment center, at least 5, and maximally 6 candidates participate together. They are, however, not competing with each other for jobs. The actual mix of the (5 or 6) persons partaking in a specific assessment center is determined only by the fact that each of those graduates has passed the preceding selection steps successfully. For each center six assessors (or 5 depending on the number of candidates) are selected randomly from a pool of about 80 senior managers. In no case can a senior manager become the assessor of a candidate he/she already has interviewed in the previous step of the *management interview*. Two exercises were used to elicit behavior from the candidates during the assessment center: a group discussion and an analysis/presentation exercise. The exercises and the end meeting take place on the same day. First six assessors observe and rate the behavior of 6 candidates in the *group discussion*, in which candidates have to come up with a common solution to a fictitious business problem in which they have conflicting

interests. During the group discussion, all (6) assessors are present. In the analysis/presentation exercise the candidate has to present an analysis, and a corresponding plan of action, for another business problem. Only two assessors observe and discuss with the candidate his/her proposals and corresponding arguments. In both exercises, assessors observe and rate the behavior of each candidate individually and separately on the dimensions thinking, interpersonal effectiveness and firmness. They do not give an end rating for the exercise and do not discuss their ratings before the end meeting. Since the assessors did not give final ratings for the group discussion and the analysis/presentation exercise, an actuarial end score was computed as the mean of all the dimension ratings given by the assessors in the exercise.

The dimension ratings given by the individual assessors in the group discussion (18 ratings in total: 6 managers rate the candidate's behavior on 3 dimensions) and the analysis/presentation exercise (6 ratings in total: 2 managers rate the candidate's behavior on 3 dimensions) have to be combined to one final OAR. In the end meeting, candidates are discussed one after the other. For each candidate, managers start with reporting their observations and evaluations of his/her performance in the group discussion. Evaluations are compared and, in case of too large a difference, discussed in terms of underlying observations. The same procedure is followed for the analysis/presentation exercise. When assessors feel they have all the required information, they individually make their final decision. Every assessor gives his/her final rating of the candidate, based on both the own ratings from the group discussion and analysis/presentation exercises, and the discussion during the end meeting. Each assessor rates the candidate as 'insufficient' (*not acceptable, reject*), 'sufficient' (*average growth expectation, suited for the job but presently not a potential for top management*), or 'good' (*high growth expectation, potential top level manager*). The category with the most ratings is taken as the final OAR. *The OAR is not communicated*, neither to the candidate nor to the manager who will become his/her 'boss'. Candidates with ratings 'sufficient or 'good*'* are invited to join the company.

Predictors and criterion

Predictors

At the end of the recruitment interview and the management interview, the 5 assessment dimensions thinking, interpersonal effectiveness, firmness, ambition and operational competence, are rated on a five-point scale, ranging from 1 ('poor'), via 2 ('insufficient'), 3 ('average'), 4 ('good') to 5 ('very good'). For the two assessment center exercises of group discussion and analysis/presentation exercise, only the first three dimensions thinking, interpersonal effectiveness, and firmness are rated. The final OAR refers to a three-point-scale, ranging from 0 ('insufficient; reject'), via 1 ('sufficient'), to 2 ('good').

Raw scores on the 9 paper and pencil test of the mental test are firstly recomputed into so-called 'factor scores' on the four factors of *numerical ability*, *analytical ability*, *verbal ability*, and *creativity*. Factor analyses showed that 9 tests indeed measures these four general and relatively independent intelligence factors. The factor scores are, secondly, transformed into a stanine normal distribution (Guilford, 1965) using the test score distributions of all academic graduates who have been tested for the company in the past 10 years (including persons who took part in selection procedures for other, non-managerial jobs). These 'Guilford 9 stanines' are, thirdly, transformed into a 5-point scale ranging from 1 ('low', stanine 1; percentiles 0-4), via 2 ('below average', stanines 2 and 3; percentiles 4-23), 3 ('average', stanines 4, 5 and 6; percentiles 23-77), 4 ('above average', stanines 8 and 9; percentiles 77-96), to 5 ('high', stanine 9; percentiles 96-100). The final test score is computed as the average of the latter 5-point scale scores across the four dimensions of mental ability. Persons with scores 1 and 2 were rejected.

Criterion

Career success was measured as salary obtained at 31 December. Data were collected from 1989 to 2002, and were corrected for collective increases as a consequence of bargaining agreements between labor unions and the company (this mostly resulted in a correction for inflation).

Analysis

We did three analyses. First, by HLM, we estimated within subjects (level-1) and between subjects (level-2) effects. At level 1, we estimated the best fitting curve for the relation between salary level and tenure, allowing linear and quadratic components only, for every subject. Level 2 analysis regressed the parameters (linear and quadratic regression weights) of the first analysis on (1) year of

entry, and total tenure (controls), and (2) ratings from the selection procedure. The final ratings of the

selection steps were not included as predictors since they are correlated with the dimensions. In a

second HLM-analysis, we included only these end ratings. Finally, we computed validities for the group

of persons with tenure of at least 12 years.

RESULTS

Predictors

Rejection percentages were 52% for the recruitment interview, 12% for the mental test, 13%

for the management interview, and 31% for the assessment center. A mere 4% of the initial job

applicants finally was hired. Table 1 presents the number of persons, means and standard deviations of

the ratings given on the dimensions assessed in the steps of the selection procedure for the group that

participated in this study. Since the assessors did not give final ratings for the group discussion and the

analysis/presentation exercise, Table 1 gives an actuarial end score computed as the mean of all the

dimension ratings given by the assessors in the exercise.

(Table 1 about here)

The group of subjects that left the company in 1997 as a consequence of the split-off of a

major business unit, had lower scores on interpersonal effectiveness as assessed in the management

interview (M=3.86, SD=.64, p=0.03) and on the dimensions interpersonal effectiveness (M=3.64,

SD=0.83, p=0.002), firmness (M=3.82, SD=0.88, p=0.01), and the mean (M=3.69, SD=0.73, p=0.01) of

the analysis/presentation exercise (see table 1 for the corresponding means and standard deviations for

the group in our study). So if there would be any effect on this split-off on the validity, it may be a slight

reduction as a consequence of a somewhat lower variance on the predictor dimensions. Furthermore, we

investigated whether there was a relation between turnover and ratings obtained in the selection

procedure by computing for all predictors the correlation with turnover. In the total group and in the

group with tenure less than 2 years there were no significant correlations. In the group with tenure

between 2 and 5 years, quitting correlated with ambition (.23; p<.03, N=89) as assessed in the

recruitment interview. In the group with tenure of 5-9 years, quitting correlated with ambition as assessed in the recruitment interview (.14, p<.05, N=209), and with numerical ability (.39, p<.001, N=219), analytical ability (.22, p<.001, N=219), creativity (.30, p<.001, N=208), and the final score (.44, p<.001, N=209) of the mental test. In the group with tenure of 9-13 years quitting correlated negatively with the assessment of thinking (-.22; p<.02, N=118) and with the mean rating (-.19, p<.04, N=118) of the group discussion. In general, it seems that persons with higher ratings on ambition in the recruitment interview are somewhat underrepresented when investigating long term validity since they tend to leave in the first half of the 13 year period. And there clearly is a strong tendency to leave for persons with higher scores on the mental test somewhere halfway the 13-year period. This leave only the lower part of the distribution of mental test scores for computing long-term validity, which (given the almost universal validity of the mental test for career success, cf. Schmidt & Hunter, 1998) could very well result in near-zero or even negative long-term validities.

We computed, for the group of selected persons, correlations between all predictors. It appeared that the selection steps are relatively independent. The group discussion and the analysis/presentation exercise are correlated, and of course the OAR is strongly related to the end scores of both assessment centers. In accordance with previous studies (e.g. Brannick, Micheals & Baker, 1989) the correlation between the same predictors across different assessment situations was lower than the correlation between different predictors within the same situation. Still, both types of correlations were not untypical for assessment center studies.

Criterion

Criterion was salary obtained in the j-th year of the career (j=1,...,13). All salaries are in the (former) Dutch currency of ƒ (1ƒ is fixed at 0,45€). The number of persons for whom criterion data were available decreased from 456 (1st year), 456 (2nd year), 453 (3rd year), 437, 415, 383, 328, 284, 211, 145, 94, 69, and finally 27 (13th year). As expected persons whose salary level stayed low or decelerated had a greater tendency to leave. In our study, the correlation between salary level and quitting the company over the period investigated was -.47 (N=456, p<.001). This may reduce the variation on the criterion. Therefore we checked the criterion for normality. Salary distributions became

more (positively) skewed to the right with advancing tenure levels, implying that increasingly more

persons earned less than average salary and a few persons obtained very high salaries (for 2, .., 9 years,

skewness was approximately 1; for 11, .., 13 years it became 1.4, 1.3, and 1.2 respectively). This effect

was not large enough to prevent subsequent (regression) analyses.

Figure 1 shows this general salary curve by depicting average salary level in the j-th year of the

career where j=1,…, 14 (in the picture we have omitted actual salary amounts because the company does

not want to publish these). Figure 1 confirms that there was no variation in starting salary, that salary

increases with tenure, that there is a (slight) tendency for a faster increase at higher tenure levels

(suggesting a quadratic curve), and that inter-individual differences in salary increase with tenure.

(Fig.1 about here)

We were not able to estimate the reliability of the criterion.

We estimated the personal salary curve by Eq. (1) for all subjects. There are however a number

of possible choices for fixing the constant $\beta(0i)$ of person's i salary curve. For instance, it could be set

equal to the initial salary (on account of company policy, starting salaries were about the same; this

was confirmed by an inspection of its distribution), or to the average salary earned in that person's

career. But since we have a quadratic model in equation (1), this choice has a non-arbitrary effect on the

general location of the curve, and therefore on the value of the linear and quadratic regression

coefficients. Since we are not interested in the prediction of the general vertical location of the salary

curve, but in its pattern of in- or decrease, it was decided to set the constant of the curve equal to the

'median salary': the salary earned at the midpoint (in terms of total tenure) of that person's career. For

instance, when a person had a total tenure of 7 years, the constant of his/her salary equation was set

equal to the salary earned at the 4th year of the career (4 being the midpoint of the series 1, 2,.., 7) but

with a total tenure of 8 years, it would be the salary earned at the $4\frac{1}{2}^{th}$ career year (which is the midpoint

of the series 1, 2,…, 8). In the random effects analysis (and consequently in the level-2 analysis), this

constant will be omitted since it is just an arbitrary vertical location of the curve.

Hierarchical linear modeling: Individual predictors

The *level-1 analysis* is an 'unconditional' analysis in the sense that level-2 predictors are not taken into consideration. The analysis consists of two steps: a 'fixed effects' analysis in which the parameters β from Eq. (1) are not allowed to vary across individuals. In the second, 'random effects' part the β are allowed to vary across individuals. The first, fixed effects analysis, showed that a model in which only mean, subject-independent linear and quadratic salary trends are incorporated, did not describe well the individual variation in salaries (figure 1 in effect illustrates this). Average median salary $\beta(00)$ was equal to ƒ 6349.4, average linear yearly increase $\beta(10)$ ('linear trend') in salary was ƒ 583.9, and average quadratic salary trend $\beta(20)$ was ƒ 40.6.

In the second, random effects analysis the degree of individual deviation from the average linear and quadratic trend was analyzed by estimating the variance components for these effects. In that way we could be certain that subjects really differed in their linear and quadratic salary trends. The estimated variances for the linear and quadratic trends were 54,431.1 and 4,430.5 respectively. The HML analysis showed these to be significantly different from zero (χ^2=68.55; p<0.000001, and χ^2=96.62, p<0.000001, respectively), so that we may conclude that persons differed in their linear and quadratic salary trends. Since the salary curve parameters are significant, they can not be omitted in describing individual salary trend. The analysis showed that total variance in salaries consisted of 16% between-subjects salary variance and 84% within-subjects salary variance. In general, within-person variation of salaries could be attributed to a linear trend for 92,3% and to a quadratic trend for 2,5%.

HLM also yields reliability estimates of the linear and quadratic regression parameters. These indicate the proportion of true or systematic between-subjects variance that can be modeled in the level-2 analysis using the individual difference variables. The weighted mean reliability of the linear salary trend $\pi(1i)$ is .91, varying from .66 to .98. The weighted mean reliability of the quadratic salary trend $\pi(2i)$ is .60, varying from .38 to .83. These estimated reliabilities suggest that it makes sense to proceed to the level-2 analysis. That is, the risk of modeling at level-2 what in fact is error is low. HLM analysis also produces an estimate of the correlation between 'true' linear salary trend $\pi(1i)$ and true quadratic salary trend $\pi(2i)$. The correlation is .32, which indicates that quadratic increase is faster for subjects with higher linear trend in salaries.

In the conditional *level-2 analysis* (table 2), year of entry, total tenure and results from the selection procedure were used as predictors of the linear and quadratic trend parameters $\pi(1i)$ and $\pi(2i)$ of the individual salary curves as estimated in the level-1 analysis. Total Multiple R for the analysis of the linear trend is .52 (adj MR²=.25, p<.001). A total of 26.8% of the variance in linear salary trends as given by the random effects variance components, could be explained in this between-subjects analysis. The controls of year of entry and total tenure accounted for 18.5% of this variance, leaving 8.3% for the predictors from the selection procedure. As table 2 shows, the linear trend $\pi(1i)$, that is the slope of the individual salary curve in time, was predicted by year of entry (control), ambition (recruitment interview), creativity and verbal ability (mental test; negative weight for verbal ability), firmness and the interaction between tenure and firmness (mean assessor score in group discussion), and interpersonal effectiveness (mean assessor score in analysis/presentation exercise).

(Table 2 about here)

Total Multiple R for the analysis of the quadratic trend $\pi(2i)$ was .32 (adj MR²=.09, p<.001). The MR² shows that of the total variance in quadratic salary trends as given by the random effects variance components, 10% could be explained in this between-subjects analysis. The controls of year of entry and total tenure accounted for 9% of this variance, leaving 1% for the predictors from the selection procedure. As table 2 shows the quadratic trend was predicted (negatively) by the interaction between tenure and operational competence as assessed in the management interview. This implies that higher rating on operational competence is associated with lower salary acceleration at higher tenure levels, so with a slower career increase.

Hierarchical linear modeling: End scores and end ratings

In the previous HML-analysis we did not include the end scores or end ratings of the selection steps since they are correlated with the dimension ratings. Therefore we repeated the HML-analysis, this time including only the end scores or ratings of the recruitment interview, mental test, management interview, and assessment center. An actuarial OAR was computed as the average of all dimension

ratings of all assessors in the group discussion and case interview. The *level-1 analysis* is the same as

in the previous analysis, only in this case other predictors are applied in the level-2 analysis to explain

variation in salary trend parameters. Explained variance in the *level-2 analysis* for the linear trend was

23% (MR=.48; adj MR²=.22, p<.001). The controls of year of entry and total tenure accounted for 19% of

this variance, leaving 4% for the predictors from the selection procedure. Year of entry (control; β=.43, SE

β=.05, p<.001), total tenure (control; β=.49, SE β=.05, p<.001), the actuarial OAR ($\triangle$R²=.03; β=.14, SE

β=.04, p<.01) and the interaction between tenure and the actuarial OAR ($\triangle$R²=.01; β=.11, SE β=.05,

p<.05) predicted the linear trend in the salary curves. So it appears that the OAR is a predictor of the

slope of the salary curve (that is of promotion speed), and that this in particular is the case at higher levels

of tenure, that is later in the career. The analysis for quadratic effects in the salary curve only yielded

tenure as (negative) predictor (β=-.26, p<.001; MR=.29; adj MR²=.08, p<.001), implying that with

increasing tenure the acceleration of the salary curve slows down.

Long-term validities for persons with tenure of at least 12 years

Finally, we limited the analysis to those persons with tenure of at least 12 years, and took final

salary level as criterion. The resulting validities, corrected for differences in starting salary, restriction in

range and predictor unreliability (using table 1 and the estimates from Jansen & Stoop, 2001; see also the

overview of OAR reliabilities from various assessment center studies in the Appendix A in Collins et al.,

2003), are presented in table 3.

(Table 3 about here)

Table 3 shows that when we try to predict only the end level of the salary curve, instead of its general

linear increase or quadratic increase, a more elaborated result is obtained. While the HLM-analysis is

confirmed (with the exception of creativity as measured by the mental test), we have additional

contributions of thinking (management interview, group discussion, analysis presentation exercise),

interpersonal effectiveness (group discussion, analysis presentation exercise), and firmness (recruitment

interview, management interview, group discussion, analysis presentation exercise). The latter findings

agree with the validities found in earlier studies in long-term assessment center validity (Bray et al., 1974; Hinrichs, 1978); but these studies had a range of 8 years. In this phase of the career, both verbal ability and analytical ability (mental test) have negative contributions. Turning to the selection steps, clearly both assessment center exercises have substantial long-term (corrected) validities.

DISCUSSION

Validity

The analyses showed that there were systematic linear and quadratic trends in the salary curves, that there were also between-person differences, and that all steps in the selection procedure were involved in the prediction of the latter. Recruitment interview, mental test, management interview, group discussion and analysis/presentation exercise all contributed to the prediction of salary growth, but with different dimensions. Since construct-validity of the dimensions was low and every selection step contributed unique dimensions to the prediction (so that dimensions and exercises more or less merge), we can as well switch from a person-based (dimensions) to a task-based (exercises or selection steps) interpretation (Schleicher et al., 2002). In general, our analyses confirmed the expectations based on Boudreau, Boswell & Judge (2001), Ferris et al. (2001), Hogan & Holland (2003), Judge et al. (1995, 2002), Robertson (2000), Seibert & Krainer (2001), and Seibert, Crant & Krainer (1999), that persuasive and initiating behaviors ('firmness' and 'ambition') and interpersonal behaviors ('interpersonal effectiveness') are important determinants of long-term management progress, that operational behavior ('operational competence') in time becomes a contra-indicator for career success in management, and that the impact of general mental ability ('thinking'), and in particular openness, seems to level off with longer tenure levels. Whereas resource-oriented, problem solving competencies determine initial performance, people-oriented competencies are more important for later performance (Russell, 2001). Firmness as measured in the group discussion predicts salary growth and its predictive power even increases with length of the career. Also ambition as assessed in the recruitment interview, and interpersonal effectiveness as measured in the analysis/presentation exercise predict career success. The interaction between tenure and operational competence as assessed in the management interview is a negative predictor of quadratic salary trend. Since when the interaction increases the weight decreases,

operational competence is a contra-indicator of (accelerated) career progress in management that becomes stronger with length of the career.

Turning from dimensions to exercises, it is confirmed that a selection interview, if properly conducted, can be a predictor of even long-term success, and that it derives it predictive power from the elicitation and assessment of motivation, initiative and drive loaded behaviors The mental test factor of creativity contributed to career success, verbal ability was a negative predictor. The latter was also found in our previous study. When the analysis is limited to the long-tenure group only, an additional negative validity of analytical ability was obtained. In this respect earlier findings of declining cognitive test validity are confirmed (Keil & Cortina, 2001). The assessment center exercises derive their predictive power from the elicitation and measurement of interpersonal, stress-resistant, and initiating behaviors. It appeared that the OAR is a good predictor of (the linear trend in) promotion speed and that it becomes a better predictor later in the career. These effects were all corrected for year of entry and total tenure. The coefficients for year of entry indicate that later cohorts had faster salary growth. Since later cohorts did not differ from earlier with respect to performance in the selection procedure, it seems that, in time, company promotion policies have somewhat changed.

The results were, in general, confirmed by separate analyses on persons with tenure of at least 12 years. The validity of the (clinical) OAR for predicting salary level over a period of at least 12 years, corrected for differences in starting salaries, was .23 (N=69, p=.03; see table 3). Corrected for differences in starting salaries, for restriction in range, and for unreliability of the predictor, it becomes .39. It can be argued that one should not correct for unreliability of the predictor since what counts is the operational validity of the selection instrument-as-used instead of its maximally attainable theoretical validity in case of perfect measurement conditions (Salgado & Anderson, 2003). On the other hand, we did not correct for unreliability of the criterion. But, salary increase is based on performance appraisal, and from meta-analyses a fairly consistent value of .52 is obtained as average reliability for job performance or training success criteria in both American and European studies (Salgado & Anderson, 2003). If we apply this value, the validity of the OAR for predicting salary level over a period of at least 12 years, corrected for differences in starting salaries, for restriction in range, and for unreliability of the criterion, becomes .49.

If we compare the validity of the (clinical) OAR (for easiness of comparison we only correct for

starting salary) to the *actuarial* validity computed as the average of all assessment center ratings, it

appears that the latter is consistently substantially larger. For persons with a tenure of at least 12 years,

the clinical validity is .23 and the actuarial validity is .49 (N=53; p=.001), and for the group with a tenure of

at least 13 years, the clinical validity is .21 (N=27, p=.15) and the actuarial validity is .45 (N=16; p=.04).

Part of this difference can be explained by the clinical validity being a point-biserial correlation with

reduced predictor variance. Still, if we correct the actuarial validities for attenuation and unreliability, large

values for the long-term predictive validity of the OAR can be obtained. For instance, correction for

unreliability of the criterion yields an actuarial OAR-validity of .68 for persons with tenure of at least 12

years. One may conclude that by using the clinical OAR for deciding on acceptance/rejection, the

assessment center results are *under*used in the present organization. It would be better practice to

compute the sum of all assessment center ratings and determine the optimal cut-off point for

acceptance/rejection on this actuarial OAR variable in such a way that the correlation between OAR and

long-term salary is maximized.

Year of entry, tenure and inter-individual differences as measured in the selection procedure

accounted for in total 34% of the between-subjects variation in salary trend. So also other, yet

unmeasured personal and situational variables are needed. We suggest that these be sought in the job

environment. For instance, although the subjects all aimed for careers in management, the fields differed,

for instance personnel, finance, commerce, or technics.

Career growth as a tournament with changing tasks

Meta-analyses show large rank-order trait consistency across persons. In the study by

Roberts & Delvecchio (2000), the trait with the largest between-persons consistency was the Big

Five-factor Extraversion. Extraversion, which maybe seen as a combination of the dimensions

interpersonal effectiveness and firmness in our study, had a mean construct-level correlation

(correlation corrected for both predictor and criterion unreliability) of .50 with the OAR in the meta-

analysis by Collins et al. (2003). So the stable factor of Extraversion may be seen as the 'driving

force' of long-term assessment center validity. In case of such a long-term between-persons stability,

a changing-tasks model best describes the dynamic pattern in validities predicting career growth

(Mitchel, 1975).

Being successful on the advancement criterion requires mastery of different combinations of the dimensions measured by the interview, mental test, and assessment center exercises at different stages in the career. That is, task requirements change with career growth and therefore other dimensions are required for mastering the new tasks and, consequently, winning the competition for better-paid jobs. The nature of managerial work requires both intellectual and social skills; in the early career intellectual abilities are decisive for career success, but after 5-9 years social skills become decisive for job effectiveness and therefore for career growth. This may be illustrated by the finding that persons leaving generally were lower in career growth, that is they were less successful in the organization at issue: a postal and telecommunications company with a technical core, which moreover was recently privatized. In such domain, intelligence factors such as numerical ability and creativity may become more important than verbal ability. Later career success however requires personal effectiveness and firmness (the factor of 'Extraversion') since at those levels a career can be compared to an arena where candidates compete for attention, support, and scarce jobs (Green & Stokey, 1983).

Interpersonal competition in such tournament yields a rank-order of persons on account of which promotion decisions are taken. Such a promotion procedure is economically efficient when monitoring of individual performances is difficult (Lazear & Rosen, 1981). A decisive factor for winning the tournament is internal reputation: relevant decision-makers know, on account of in particular informal networking contacts (Gersick, Bartunek & Dutton, 2000; Seibert, Kraimer & Liden, 2001) that a candidate has been effective a lower job levels. Since we are used to experimental or survey studies in unrestricted groups, we tend to forget that organizations represent very selective situations. At higher organizational levels, there exist a severe 'selection pressure' (Pinker, 2002) in the sense that resources that are necessary for survival (that is for management development) are scarce; for instance informal career help for building personal networks, and entry to sources of advice, support and information (Sturges, Guest & Conway, 2002). Both less management attention and less management positions at higher hierarchical levels result in a 'struggle for promotion'. And since only specific, interpersonal and 'firm', behaviors are good for survival that is for promotion to a higher level, we tend to find at higher level persons who are selected for those characteristics. This also results in

a strong restriction in range with respect to these characteristics.

Limitations of the present study and implications for further studies

There was a strong tendency to leave for persons with higher scores on the mental test somewhere halfway the 13-year period. They were lower in salary growth This leave only the lower part of the distribution of mental test scores for computing long-term validity, which may explain the negative validities obtained. So the negative validities of (some) mental test factors could be explained by turnover. The degree of range restriction (average ratio of restricted standard deviation to unrestricted standard deviation) was .93 in the average for the mental test. This is a rather infrequent value according to the empirical distributions for cognitive ability composed by Alexander et al. (1989; these authors obtained a median value of .68), or by Salgado & Anderson (average restriction ratio for the Low Countries was .50). It seems therefore that the mental test tends to be *under*used in the selection procedure. Note however that this would leave more room for the 'predictive ability' of the mental test to manifest itself, if present, in our study.

The control of year of tenure revealed that there were clear cohort effects with respect to promotion. In general it appeared that later groups were promoted faster than groups that were recruited earlier. Although analyses on the total group and the long-term group yielded comparable results, and validity results were corrected for the effect of year of entry, it cannot be excluded that career progress of the later cohort will, for instance on account of ceiling effects, be associated with a different validity profile in time.

In what way are our findings restricted to the organization at issue? Since career progress is the outcome of both general developmental processes and organization and even nation-specific promotion policies (cf. the differences in this respect between such European countries as France and the Netherlands, Rousseau & Tinsley, 1997), long-term validities will be affected by situational factors. On the one hand, in open societies with relatively unconstrained local and international competition for both goods/services and top-level managers, variation in criteria for effective executives will tend to decrease. On the other hand, cultural differences between even adjacent countries will stay, which underlines the requirement of managers to adapt for local effectiveness. But, the latter requires (apart from language

fluency) high-order faculties such as empathy, flexibility and social effectiveness, which we know from

research in expatriate success factors, and which by definition are global.

REFERENCES

Alexander, R.A., Carson, K.P., Alliger, G.M., & Cronshaw, S.F. (1989). Empirical distributions of range restricted SDx in validity studies. *Journal of Applied Psychology*, *74*, 253-258.

Anstey, A. (1977). A 30-year follow-up of the CSSB procedure, with lessons for the future. *Journal of Occupational Psychology*, *50*, 149-159.

Barrett, G.V., R.A. Alexander, & D. Doverspike. (1992). The implications for personnel selection of apparent declines in predictive validities with time: A critique of Hulin, Henry, and Noon. *Personnel Psychology*, *45*, 601-617.

Boudreau, J.W., Boswell, W.R., & Judge, T.A. (2001). Effects of personality on executive career success in the United States and Europe. *Journal of Vocational Behavior*, *58* (No.1), 53-81.

Brannick, M.T., Micheals, Ch.E., & Baker, D.P. (1989). Construct validity of in-basket scores. *Journal of Applied Psychology*, *74*, 957-963.

Bray, D. W., Campbell, R.J., & Grant, D.L. (1974). *Formative years in business: A long-term A.T.&T. study of managerial lives*. New York: Wiley.

Collins, J.M., Schmidt, F.L, Sanchez-Ku, M., Thomas, L., McDaniel, M.A., & Le. H. (2003). Can basic individual differences shed light on the construct meaning of assessment center dimensions? *International Journal of Selection and Assessment*, *11* (No.1, March), 17-29.

Deadrick, ,D.L., Bennett, N., & Russell, C.J. (1997). Using hierarchical linear modeling to examine dynamic performance criteria over time. *Journal of Management*, *23* (No.6), 745-757.

Ferris, G.R., Witt, L.A., & Hochwarter, W.A. (2001). Interaction of social skill and general mental ability

on job performance and salary. *Journal of Applied Psychology*, *86* (No.6), 1075-1082.

Fiedler, F.E. & House, R.J. (1994). Leadership theory and research: A report of progress. In: C.L. Cooper & I.T. Robertson (Eds.). *Key reviews in managerial psychology*. New York: Wiley, 97-116.

Gaugler, B.B., Rosenthal, D.B., Thornton III, G.C., & Bentson, C. (1987). Meta-analysis of assessment centre validity. *Journal of Applied Psychology*, *72*, 493-511.

Gersick, C.J.G., Bartunek, J.M., & Dutton, J.E. (2000). Learning from academia: The importance of relationships in professional life. *Academy of Management Journal*, *43* (No.6, December), 1026-1044.

Green, J.R., & Stokey, N.L. (1983). A comparison of tournaments and contracts. *Journal of Political Economy*, *91*, 349-364.

Guilford, J.P. (1965). *Fundamental statistics in psychology and education*. New York: McGraw-Hill.

Guilford, J.P. (1967). *The nature of human intelligence*. New York.

Hinrichs, J.R. (1978). An eight-year follow-up of a management assessment center. *Journal of Applied Psychology*, *63* (No.5), 596-601.

Hofmann, D.A. (1997). An overview of the logic and rationale of hierarchical linear modeling. *Journal of Management*, *23* (No.6), 723-744.

Hogan, R., Curphy, G.J., & Hogan, J. (1994). Wat we know about leadership. Effectiveness and personality. *American Psychologist*, *49* (No.6, June), 493-504.

Hogan, J. & Holland, B. (2003). Using theory to evaluate personality and job-performance relations: A

socioanalytic perspective. *Journal of Applied Psychology*, 88 (No.1), 100-112.

Huck, J.R. (1977). The research base. In: J.L. Moses & W.C. Byham (Eds.). *Applying the assessment center method*. New York: Pergamon Press, 261-291.

Hulin, Ch.L., R.A. Henry, & S.L. Noon. (1990). Adding a dimension: Time as a factor in the generalizability of predictive relationships. *Psychological Bulletin*, *107* (No.3), 328-340.

Hunter, J.E., & Hunter, R.F. (1984). Validity and utility of some alternative predictors of job performance. *Psychological Bulletin*, *96*, 72-98.

Jansen, P.G.W., Stoop, L.A.M. (2001). The dynamics of assessment center validity: Results of a seven year study. *Journal of Applied Psychology*, *86* (No.4, Aug), 741-753.

Judge, T.A., Cable, D.M., Boudreau, J.W., & Bretz, R.D. jr. (1995). An empirical investigation of the predictors of executive career success. *Personnel Psychology*, *48* (No.3), 485-519.

Judge, T.A. , Bono, J.E., Ilies, R., & Gerhardt, M.W. (2002). Personality and leadership: A qualitative and quantitative review. *Journal of Applied Psychology*, *87* (No.4), 765-780.

Keil, C.T., & Cortina, J.M. (2001). Degradation of validity over time: A test and extension of Ackerman's model. *Psychological Bulletin*, *127* (No.5), 673-697.

Lazear, E.P., & Rosen, S. (1981). Rank-order tournaments as optimum labor contracts. *Journal of Political Economy*, *89*, 841-864.

McEvoy, G.M., & Beatty, R.W. (1989). Assessment centres and subordinate appraisal of managers: A seven year examination of predictive validity. *Personnel Psychology*, *42*, 37-52.

Mitchel, J.O. (1975). Assessment center validity: A longitudinal study. *Journal of Applied Psychology*, *60* (No.3), 573-579.

Moses, J.L. (1971). Assessment center performance and management progress. *Paper presented at the 79ᵗʰ annual meeting of the American Psychological Society*, Washington, D.C..

Pinker, S. (2002). *The blank slate. The modern denial of human nature*. Harmondsworth: Viking Penguin.

Roberts, B.W., & Delvecchio, W.F. (2000). The rank-order consistency of personality traits from childhood to old age: A quantitative review of longitudinal studies. *Psychological Bulleting*, *126* (No.1), 3-25.

Robertson, I.T. (2000). Conscietiousness and managerial performance. *Journal of Occupational and Organizational Psychology*, *73*, 171-180.

Rousseau, D.N. & Tinsley, C. (1997). Human resources are local: Society and social contracts in a global economy. In: N. Anderson & P. Herriot (Eds.). *International handbook of selection & assessment*. New York: Wiley, 39-61.

Russell, C.J. (2001). A longitidinal study of top-level executive performance. *Journal of Applied Psychology*, *86* (No.4), 560-573.

Schleicher, D.J., Day, D.V., Mayes, B.T., & Riggio, R.E. (2002). A new frame for frame-of-reference training: Enhancing the construct validity of assessment centers. *Journal of Applied Psychology*, *87* (No.4), 735-746.

Salgado, J.F., & Anderson, N. (2003). Validity generalization of GMA tests across countries in the

European Community. *European Journal of Work and Organizational Psychology*, *12* (No.1), 1-17.

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* (No.2), 262-274.

Seibert, S.E., Crant, J.M., & Kraimer, M.L.. (1999). Proactive personality and career success. *Journal of Applied Psychology*, *84* (No.3), 416-427.

Seibert, S.E. , & Krainer, M.L. (2001). The five-factor model of personality and career success. *Journal of Vocational Behavior*, *58* (No.1), 1-21.

Seibert, S.E. , Kraimer, M.L., & Liden, R.C. (2001). A social capital theory of career success. *Academy of Management Journal*, *44* (No.2, April), 219-237.

Sturges, J., Guest, D., Conway, N., & Mackenzie Davie, K. (2002). A longitudinal study of the relationship between career management and organizational commitment among young graduates in the first ten years of work. *Journal of Organizational Behavior*, *23* (No.6), 731-748.

Sturman, M.C. & Trevor, C.O. (2001). The implications of linking the dynamic performance and turnover literatures. *Journal of Applied Psychology*, *86* (No.4), 684-696.

Thornton, G.C. III, & Byham, W.C. (1982). *Assessment centers and managerial performance*. New York: Academic Press.

Tziner, A., Ronen, S., & Hacohen, D. (1993). A four-year validation study of an assessment center in a financial corporation. *Journal of Organizational Behavior*, *14*, 225-237.

Authors Note

Paul G.W. Jansen, Department of Management & Organization Studies, Faculty of Economics and Business Administration, Vrije Universiteit Amsterdam, the Netherlands;

Table 1.

Mean, and standard deviations for the predictors for those managers for whom criterion data were available

Predictors	M	SD
RECRUITMENT INTERVIEW		
Thinking (*)	4.08	.42
Interpersonal effectiveness (*)	4.08	.44
Firmness (*)	4.14	.41
Ambition (*)	4.09	.45
Operational Competence (*)	4.22	.40
Final rating (*)	4.12	.32
MENTAL TEST		
Numerical ability (**)	3.03	.84
Analytical ability (**)	3.23	.75
Verbal ability (**)	3.02	.67
creativity (**)	3.36	.75
Final test score (**)	3.11	.85
MANAGEMENT INTERVIEW		
Thinking (*)	4.01	.66
Interpersonal effectiveness (*)	3.99	.71
Firmness (*)	3.94	.68
Ambition (*)	3.84	.74
Operational Competence (*)	3.89	.76
Final rating (*)	3.96	.53
GROUP DISCUSSION		
Thinking (*)	3.76	.48
Interpersonal effectiveness (*)	3.60	.60
Firmness (*)	3.72	.58
mean of assessors (*)	3.70	.49
ANALYSIS/PRESENTATION EXERCISE		
Thinking (*)	3.70	.84
Interpersonal effectiveness (*)	3.84	.77
Firmness (*)	4.00	.77
mean of assessors (*)	3.85	.67
(clinical) OVERALL ASSESSMENT RATING (OAR) (***)	1.37	.48

Note. N varies between 409 and 456

*: Score range: 1 ('poor'), 2 ('insufficient'), 3 ('average'), 4 ('good'), 5 ('very good').

**: Score range: 1 ('low', stanine 1; percentiles 0-4), 2 ('below average', stanines 2 and 3; percentiles 4-

23), 3 ('average', stanines 4, 5 and 6; percentiles 23-77), 4 ('above average', stanines 7 and 8; percentiles 77-96), to 5 ('high', stanine 9; percentiles 96-100).

***: Score range: 0 ('insufficient'; reject), 1('sufficient'), 2 ('good').

.

Table 2

Hierarchical linear modeling analysis: Level-2 model for salary. Regression of random effect weights from level-1 analysis on starting year and total tenure (controls), and on selection dimensions (stepwise: only dimensions with significant regression coefficients are reported; standardized regression coefficients as obtained in the final equation)

Fixed effects	β	SE β	ΔR²
	Model for linear trend in salary: $\pi(1i)$		
(Constant)	.00	.04	
(1) Control:			.185
year of entry	.50***	.06	
total tenure	.47***	.05	
(2) selection procedure:			
firmness (GD)	.14**	.05	.027
ambition (RI)	.15**	.05	.017
tenure x firmness (GD)	.10*	.04	.013
creativity (MT)	.12*	.05	.008
verbal intelligence (MT)	-. 11*	.05	.010
interpersonal effectiveness (AP)	.09*	.05	.008
	Model for quadratic trend in salary: $\pi(2i)$		
(Constant)	-.01	.05	
(1) Control:			.09
year of entry	.13*	.06	
total tenure	-.23**	.06	
(2) selection procedure:			.01
tenure x operational competence (MI)	-.10*	.05	

Note. N=378. RI: recruitment interview; MT: mental test; MI: management interview; GD: group discussion (mean assessor score); AP: analysis/presentation exercise (mean assessor score).

ΔR² is the incremental variance explained by the corresponding predictor(s) in addition to the variance explained by predictors from the previous steps. Total Multiple R² for the linear trend is .27, and for the quadratic trend is .10.

*p<.05, two-tailed. **p<.01, two-tailed. ***p<.001, two-tailed.

Table 3.

Correlations between predictors and final salary (corrected for starting salary) for persons with tenure of at least 12 years. Only in case of significance (p≤05) correlations are presented corrected for restriction in range and for predictor unreliability

Predictors	validity	Sig. (1-tailed)	Validity corrected for	
			restriction in range	restriction in range and predictor unreliability
RECRUITMENT INTERVIEW				
Thinking	.03	.40		
Interpersonal effectiveness	.01	.47		
Firmness	.22	.05	.40	*
Ambition	.18	.09		
Operational Competence	.00	.49		
MENTAL TEST				
numerical ability	-.01	.48		
analytical ability	-.23	.03	-.25	-.28
verbal ability	-.21	.04	-.24	-.27
creativity	.13	.14		
MANAGEMENT INTERVIEW				
Thinking	.22	.04	.25	*
Interpersonal effectiveness	.11	.20		
Firmness	.22	.04	.25	*
Ambition	.20	.06		
Operational Competence	.18	.09		
GROUP DISCUSSION				
Thinking	.41	.001	.45	.56
Interpersonal effectiveness	.34	.005	.39	.47
Firmness	.43	.001	.50	.62
ANALYSIS/PRESENTATION EXERCISE				
Thinking	.31	.009	.39	*
Interpersonal effectiveness	.37	.003	.44	*
Firmness	.27	.020	.34	*
Clinical Overall Assessment Rating	.23	.031	.35	.40

Note. N varies between 53 and 69.

*: It was not possible to compute reliabilities for both interviews and the analysis/presentation exercise.
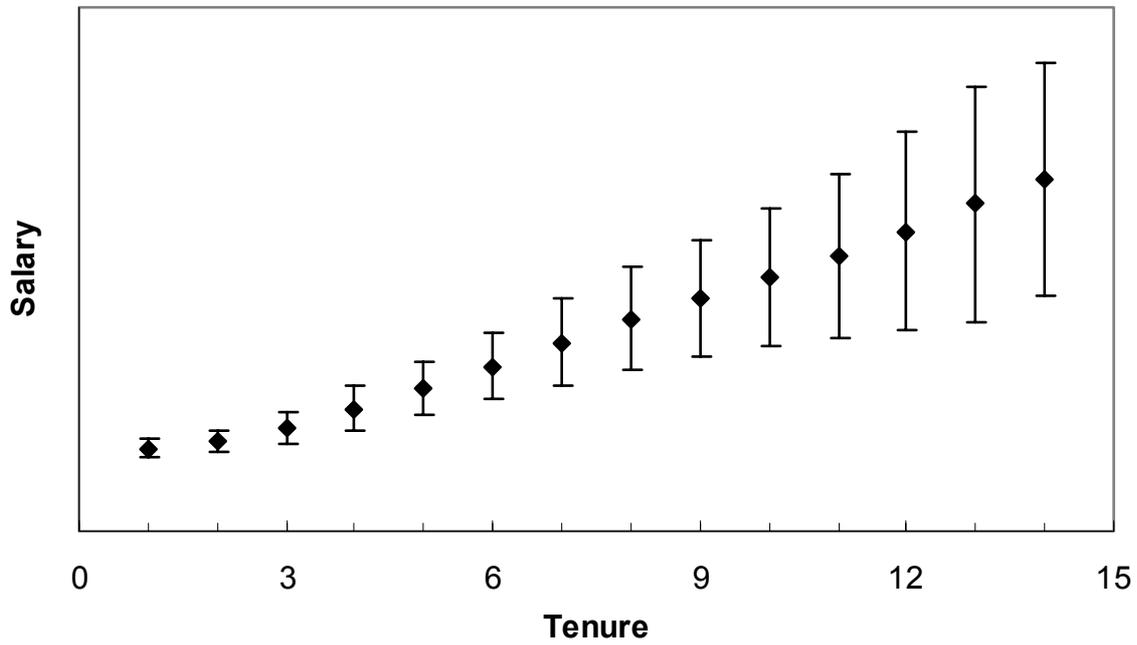
Figure 1. Range of individual salary curves.
Absciss: salary as obtained in the j-st year of the career, j=1,.., 14).
Ordinate: salary level (actual figures can not be published)
♦: average salary. T : maximum salary. ⊥: minimum salary.