

## PREDICTING TYPICAL AND MAXIMUM PERFORMANCE WITH MEASURES OF MOTIVATION AND ABILITIES

Ute-Christine KLEHE<sup>(1)</sup> & Gary LATHAM<sup>(2)</sup>  
(1) *University of Amsterdam, the Netherlands &*  
(2) *University of Toronto, Canada*

The current study integrated the literature on selection tests of typical versus maximum performance (Cronbach, 1960) with the literature on job performance under typical and maximum performance conditions (Sackett, Zedeck, & Fogli, 1988). Tests of maximum performance (i.e., measures of task-related knowledge, skills, and abilities) loaded onto a different factor than tests of typical performance (i.e., measures of task-related motivation). Nevertheless, these two factors were moderately correlated ( $r = .44$ ). Further, both task-related ability and motivation predicted typical performance. Maximum performance was predicted only by ability.

Cronbach (1949; 1960) classified personnel selection tests into two broad categories, tests of maximum and tests of typical performance. The former term designates ability tests such as measures of cognitive ability or tests of a particular knowledge or skill. The distinguishing feature of maximum performance tests is that they seek to assess how much or how well people can perform at their best. Hence, candidates are encouraged to do well in order to earn the best score they can. In contrast, “tests of typical performance are used to investigate not what the person can do but what he does” (Cronbach, 1960, p. 31) and “the test of a suitable employee is whether she maintains that courtesy in her daily work even when she is not ‘on her best behaviour’”. Tests of typical performance therefore do not assess a candidate’s ability or what that person *can* do under maximum circumstances, rather they assess how the ability of a job candidate is evident on a day-to-day basis. In short, these tests assess what the candidate *chooses* to do. Hence they assess a candidate’s motivation rather than his or her ability.

Applying Cronbach’s classification from the selection domain to actual job performance, Sackett, Zedeck, and Fogli (1988; see also Sackett, 2007) argued that typical job performance occurs when people are not aware that they are being evaluated on the job, when they are not instructed to do their ‘very best’, and when their job performance is assessed over an extended

---

Ute-Christine Klehe, Programmagroep Arbeids- & Organisatiepsychologie, University of Amsterdam, the Netherlands. Gary P. Latham, Rotman School of Management, University of Toronto, Canada.

Correspondence concerning this article should be addressed to Ute-Christine Klehe, University of Amsterdam, Programmagroep A&O Psychologie, Roetersstraat 15, NL-1018 WB Amsterdam, the Netherlands. E-mail: [u.klehe@uva.nl](mailto:u.klehe@uva.nl)

period of time. In contrast to typical job performance, maximum job performance occurs when people know that their performance is being evaluated, when they receive instructions to exert great effort, and when the duration of assessment is short enough to enable performers to remain focused on the task. Sackett et al. (1988; see also Sackett, 2007) acknowledged that in practice, typical and maximum job performance represent a continuum rather than a dichotomy, thus rendering any comparison between typical and maximum performance relative.

Tests of maximum versus typical performance on the predictor side and maximum versus typical job performance on the criterion side are obviously related. Yet, very few empirical studies have been conducted that take into account these two domains. Instead, past studies on the prediction of typical versus maximum job performance have usually focused on selection tests of only maximum (e.g., DuBois, Sackett, Zedeck, & Fogli, 1993) or only typical performance (e.g., ForsterLee, 2007; Lim & Ployhart, 2004; Ployhart, Lim, & Chan, 2001). Exceptions to this statement have either been studies conducted in a laboratory setting (Klehe & Anderson, 2007b) or studies that have used criteria of typical versus maximum performance that lacked comparability (Marcus, Goffin, Johnston, & Rothstein, 2007; Sackett, 2007). Thus the purpose of the current study was to provide an externally valid investigation of the prediction of typical versus maximum job performance on the same criterion (Sackett et al., 1988) with a set of task-related selection tests of typical versus maximum performance on the predictor side (Cronbach, 1960).

### Typical and maximum job performance

Job performance is a function of an individual's job-related ability and motivation (Locke, Mento, & Katcher, 1978; Maier, 1955). Job-related ability comprises requisite declarative and procedural knowledge and skills (Campbell, 1990; Campbell, Gasser, & Oswald, 1996; Campbell, McCloy, Oppler, & Sager, 1993; Hunter, 1983; Motowidlo, Borman, & Schmit, 1997), which in turn are a function of other factors such as performers' general mental ability. Motivation, is defined as a performer's choice, effort, and persistence (Campbell, 1990). The argument underlying typical versus maximum performance on the criterion side is that in job situations requiring maximum performance, motivation is constrained to be high (DuBois et al., 1993; Sackett, 2007; Sackett et al., 1988). This is because people know that they are being monitored for a brief period of time, and hence they accept instructions to exert ample effort. They primarily, if not solely, focus on the task at hand. Hence, while persistence is not an issue, "Unless one is inviting disci-

plinary action, one has little choice but to expend effort on the task in question ... (and) accepting these instructions (to focus full attention on optimal performance) leads to a high level of effort" (DuBois et al., 1993, p. 206). Consequently, Sackett et al. (1988) concluded that differences between people on a measure of maximum job performance reflect primarily individual differences in ability.

Typical job performance reflects choice, effort and persistence in a relatively unconstrained situation. It reflects what a person 'chooses to do' in drawing upon ability to execute a task. As Janz (1989, p. 164) stated, "Maximum performance focuses on competencies, whereas typical performance focuses on choices" – even though such choices are affected by an individual's ability.

The distinction between typical and maximum job performance has implications for work settings (e.g., Dewberry, 2001; Smith-Jentsch, Salas, & Brannick, 2001; Thorsteinson & Balzer, 1999), particularly personnel selection (e.g., Ackerman & Humphreys, 1990; Arvey & Murphy, 1998; Borman, 1991; Klehe & Anderson, 2005). Guion (1991) proposed that the relatively low correlation between measures of typical and maximum performance on the same task explains in part the low criterion-related validity coefficients found for many predictors of job performance. Similarly, Campbell (1990) argued that hiring people on the basis of a predictor test of maximum performance could be one cause for the weak relationship that is subsequently found between the results of that selection decision and a person's typical performance on the job. Boudreau (1991) pointed out that a mismatch between a predictor (e.g., of maximum performance) and a criterion measure (e.g., typical performance) adversely affects the results of utility analyses. Consequently, both researchers and practitioners need to make explicit the aspect of performance they wish to predict, and choose the appropriate predictor to assess it (Guion, 1998).

Empirical research on typical versus maximum performance has been relatively non-existent, leading Sackett (2007, p. 185) to describe "the typical-maximum performance work as my biggest disappointment, in the sense that I see it as under-researched relative to the usefulness of the distinction". However, research activity in this area has increased since the turn of the millennium. A special issue of *Human Performance* (Klehe, Anderson, & Viswesvaran, 2007) addressed the topic of typical versus maximum performance from diverse and novel perspectives (e.g., Barnes & Morgeson, 2007; Mangos, Steele-Johnson, LaHuis, & White, 2007; Mesmer-Magnus & Viswesvaran, 2007; Smith-Jentsch, 2007; Witt & Spitzmüller, 2007). Different studies have examined the distinction between typical and maximum performance on sensory motor, verbal, interpersonal, and administrative tasks in field and laboratory settings (e.g., DuBois et al., 1993;

ForsterLee, 2007; Klehe & Anderson, 2007b; Marcus et al., 2007; Ployhart et al., 2001; Sackett et al., 1988). Most of these studies support Sackett et al.'s (1988) hypotheses (yet see Klehe, Anderson, & Hoefnagels, 2007 for less supportive results). Nevertheless, Sackett (2007, p. 181) concluded, after reading all published studies available to date that: "[...] an appropriate unconfounded comparison would require measurement using the same modality [...], the same level of specificity [...] at the same point in time [...], and reliable measurement of both constructs. [...] studies in field settings generally do not meet these conditions [...] (and) the 'performance over the long term' aspect of typical performance is not amenable to study in short-duration lab studies. Thus, my assessment is that relatively little research has attempted to directly address the question of the magnitude of the relationship between typical and maximum performance".

A substantial number of studies in this domain are laboratory experiments (e.g., Klehe & Anderson, 2007a, 2007b; Klehe, Anderson, & Hoefnagels, 2007; Mangos et al., 2007; Smith-Jentsch, 2007). This arguably limits the generalizability of conclusions drawn to typical work performance situations. Of the field studies on typical versus maximum performance, most failed to present comparable assessments of maximum performance. Maximum performance was often measured on a construct different from that used to measure typical performance (e.g., ForsterLee, 2007; Witt & Spitzmüller, 2007). Others measured maximum performance in an assessment centre, and in doing so, presented a different type of situation than that where typical performance was measured (e.g., Ployhart et al., 2001). Still others assessed maximum performance at a different point in time than the measurement of typical performance (e.g., Marcus et al., 2007). Finally, several studies simply derived their measure of maximum performance from the maximum value achieved under typical performance conditions (e.g., Barnes & Morgeson, 2007). Finally, all but two published studies (Klehe, Anderson, & Hoefnagels, 2007; Klehe & Latham, 2006) failed to include a manipulation check in order to ensure the validity of a comparison between measures of typical versus maximum performance.

Consequently, the current study used an extended dataset from an earlier study (Klehe & Latham, 2006) to address and test the typical-maximum performance distinction in a real-life setting explicitly designed to meet all of Sackett's (2007; Sackett et al., 1988) criteria described above. Klehe and Latham (2006) examined the empirical differences between typical and maximum performance in order to infer the constructs underlying two predictors, the patterned behaviour description interview (PBDI) and the situational interview (SI). Both the PBDI and SI predicted typical performance significantly better than maximum performance, and both methods accounted for incremental validity in the prediction of typical performance after including

maximum performance in the regression equation.

The present study differs from Klehe and Latham's in that different hypotheses were tested. Second, the focus of the present study was on performance criteria rather than predictors. The present study examined the variance accounted for in typical versus maximum performance criteria by different measures of motivation and ability. Third, Klehe and Latham used correlational and regression analyses to test their hypotheses. The present study proposed and tested a model of the interdependencies of the different measures included in this study using structured equation modeling. This allowed for a more defined and complete picture of the data than that done by Klehe and Latham. The model is shown in Figure 1. Fourth, Klehe and Latham reported data collected from 79 interviewees. The present study analysed data on the typical and maximum performance of 167 people. Also, the present study analysed data on task related knowledge, skill, and ability not reported in the earlier paper.

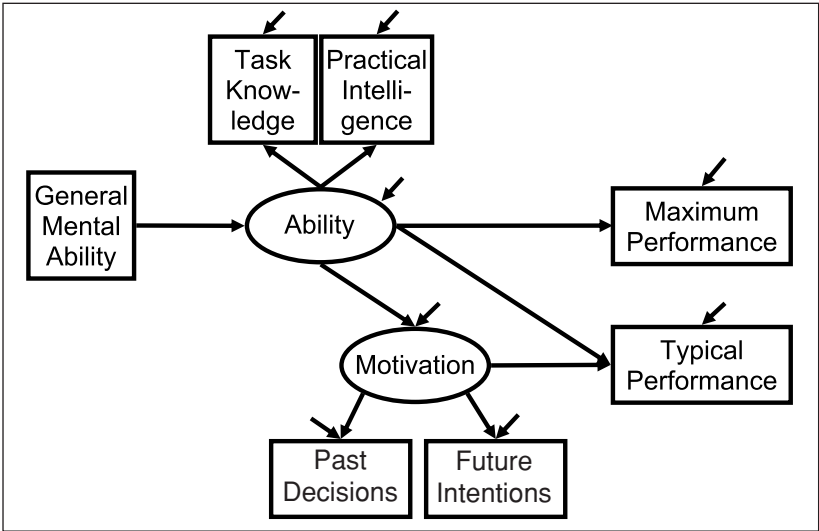


Figure 1

*The proposed model linking the predictors ability and motivation with the criteria of maximum and typical performance*

The value of the present study is that it combined the typical-maximum performance literature with the selection literature on this topic while ensuring that the data that were analysed had both external and internal validity. High external validity was ensured by the real-life setting in which the data were collected. High internal validity was ensured by the strictly task-relat-

ed nature of the predictors used to assess typical and maximum performance. In addition, there was a high level of experimental control, as indicated by diverse manipulation checks. This ruled out alternative explanations for the results. Figure 1 presents an overview of the relationships proposed in the following section.

### Hypotheses

Consistent with Cronbach's (1960) distinction between selection tests of maximum and tests of typical performance the following hypotheses were tested.

*Hypothesis 1. Selection tests that assess ability in the form of task related declarative knowledge and procedural skills load on a different factor than do selection tests that assess motivation in the form of task related choices (past decisions, future intentions)*

Candidates' choice of "what they do" is influenced by their task-related knowledge and skills in a particular domain. Candidates are less likely to make a particular choice if they do not believe their choice will lead them to their respective goals in the situation at hand (Bandura, 1986; 1997). In other words, while task-related ability cannot ensure high motivation in the respective direction, enactive mastery may well affect motivation (e.g., Bandura, 1997; Kanfer & Ackerman, 1989a, 1989b; Kanfer, Ackerman, Murtha, Dugale, & Nelson, 1994).

*Hypothesis 2. Motivation is influenced by task related ability (knowledge and skills)*

Applying Cronbach's (1960) distinction between predictive tests of typical and maximum performance to the criterion side, Sackett et al. (1988) argued that measures of motivation will not predict performance under maximum performance conditions. If a relationship between motivation and maximum performance exists, it will likely only be due to the shared effect of task related ability on both such measures. Hypotheses 2 and 4 suggest that motivation may partially mediate the relationship between ability and typical performance.

*Hypothesis 3. Maximum performance is predicted by task-related ability (task-related knowledge and skills)*

*Hypothesis 4. Typical performance is predicted by both motivation and task-related ability (task-related knowledge and skills)*

That maximum performance work settings constrain employees' motivation to be high (Sackett, 2007; Sackett et al., 1988) bears two implications: First, maximum performance situations reduce the variance of performers' motivation. Second, maximum performance work settings increase a performer's level of performance.

While arguably obvious, both assumptions have yet to be sufficiently tested. No study to date that has addressed the influence of a work setting on the variance of performance. Moreover, the results have been mixed regarding the influence of a work setting on the level of performance. Supporting evidence has been found on performers' motivation in two laboratory experiments (Klehe & Anderson, 2007a; 2007b). Support for the hypothesis on level of performance has been less conclusive. The results of two laboratory experiments (Klehe & Anderson, 2007b; Mangos et al., 2007) showed that participants performed better under maximum than under typical performance conditions. But, another laboratory experiment did not obtain this result (Klehe, Anderson, & Hoefnagels, 2007). To date, no field study has empirically addressed mean differences in typical versus maximum performance. Consequently, the fifth hypothesis of the current study addressed the question of differences in variance and level of performance between typical and maximum performance situations.

*Hypothesis 5. Performance in a maximum performance situation has (a) significantly lower variance and (b) significantly higher average than performance in a typical performance situation*

## Method

### *Setting and sample*

Students ( $n = 167$ ) at a large Canadian University provided peer evaluations of their cohorts' typical and maximum teamplaying performance during their first term in an MBA program. Participants ( $n = 120$  to 167) responded to multiple ability-measures. Participants ( $n = 79$ ) indicated their teamplaying motivation in the form of choices made in the past and inten-

tions regarding how to respond to likely future scenarios.<sup>1</sup> Participants' average age was 28 years ( $SD = 3.5$ ) with 4.7 years of work-experience ( $SD = 3$ ).

This sample of participants was used for three reasons: (1) The performance criterion was teamplaying behaviour, a core characteristic of managerial careers (Allred, Snow, & Miles, 1996; Brodbeck, Frese, Akerblom, Audia, Bakacsi, Bendova et al., 2000); (2) The predictors were administered within the same time period. This ensured that the duration between the assessment of the predictors and the criterion was relatively constant across participants; (3) The MBA program enabled a parallel assessment of typical and maximum performance in a real-life setting, as well as a manipulation check regarding these two measures.

## *Measures*

### *Criteria*

Both the typical and maximum performance criteria assessed teamplaying behaviour in an MBA program. Both were assessed through anonymous peer-appraisals on the same behavioural observation scales (BOS; Latham & Wexley, 1977; 1994) toward the end of the students' first term in the MBA program. Peer-appraisals were used because peers have many opportunities to observe each other during their work (Love, 1981; Mumford, 1983). BOS were used because they are content-valid, correlate with hard criteria, and have high inter-observer reliability (e.g., Latham & Wexley, 1977; 1994). BOS are associated with high clarity, user acceptance (Tziner & Kopelman, 1988) and ratee satisfaction (Tziner, Joanis, & Murphy, 2000). BOS also lead to specific, observable goals as well as positive ratee perceptions of these goals (Tziner et al., 2000). In sum, BOS have been associated with both high subsequent performance and satisfaction with the appraisal process (Tziner et al., 2000; Tziner, Kopelman, & Livneh, 1993). In this sense, BOS appear relatively free of the major inaccuracy- and bias-effects that afflict other forms of behaviour-based rating formats.

---

<sup>1</sup>Two primary reasons accounted for the differing sample sizes between measures: (a) In the letter inviting incoming MBA-students to participate in the interviews, students were asked to contact the author via her business-school-e-mail-address. Unfortunately, the business-school switched its e-mail systems at the time of the interviews. This caused unexpected difficulties and an unknown percentage of incoming as well as outgoing e-mails never reached their destination without causing any error messages; (b) interview-data were collected in September 2001. The attack on September 11, 2001 disrupted the collection of interview-data, leading to a loss of about forty interviews that had been scheduled for the remaining week. We deemed it neither proper nor possible to obtain further, comparable interview-data in the subsequent weeks, both due to the shock that Sept. 11 caused within the business school where the data were collected and due to participants' increasing experience with the school which rendered interviews non-comparable.



The BOS used for the assessments of both typical and maximum performance were developed by Sue-Chan and Latham (2004) to assess teamplaying in the same MBA program that was used in the current study. The scale consists of 14 behavioural items (e.g., "Coordinates upcoming work with group members who are involved with the work" and "Knows content of aspects of the project completed by other group members"), each followed by a 5-point Likert-type scale ranging from 0 (*almost never*) to 4 (*almost always*). Sue-Chan and Latham (2004) found that scores on the BOS correlate significantly with students' grade point average in the MBA program ( $r = .61, p < .05$ ). The current study used 13 of the 14 items. The item, "Knows how to use software necessary for completing assignments", was excluded, as it was not relevant to teamplaying performance.

As previously noted, typical versus maximum performance represents a continuum. Prior research has employed relatively strong manipulations, with assessments of maximum performance lasting only a few minutes (e.g., ForsterLee, 2007; Klehe & Anderson, 2007a, 2007b; Klehe, Anderson, & Hoefnagels, 2007; Mangos et al., 2007; Sackett et al., 1988; Smith-Jentsch et al., 2001), hours (e.g., Kirk & Brown, 2003; Ones & Viswesvaran, 2007), or at most a couple of days (e.g., Marcus et al., 2007; Ployhart et al., 2001). Such short observation time-periods, especially when assessing socio-psychological variables such as teamplaying, are open to at least two threats to validity when compared to longer observation periods (Cascio, 1998). First, different dimensions of performance may arise when performance in the typical performance condition includes tasks that are not relevant during a short-term maximum performance assessment (e.g., preparing for group-meetings). This lowers the validity with which one can conceptualise tasks during the typical and the maximum performance periods as they are not comparable. Second, the assessment of maximum performance is based on far fewer observations than the assessment of typical performance. This in turn threatens the comparability of reliability estimates between the two criteria (Klehe & Anderson, 2005). Equivalent reliability is required when comparing typical with maximum performance (Sackett et al., 1988).

Consequently, the current study employed a weaker manipulation in order to increase the reliability of the assessments of maximum performance. Rather than assess maximum performance over 2 days, as was done by Ployhart et al. (2001), we assessed maximum performance over a 5-day period. This was done with the knowledge that a weak manipulation, 5-days, as opposed to a strong manipulation, a few minutes (e.g., Klehe & Anderson, 2007b), might produce relatively small statistical effects. However, as Prentice and Miller (1992) pointed out, the inferences drawn from a weak manipulation are more likely to generalise to other settings than are those made on the basis of a strong manipulation.

Consequently, maximum performance in the present study was defined as the arithmetic mean of scores given anonymously to a participant by peers ( $n = 4-5$ ) during a 1-week/5-day team project. This project accounted for 25% of the students' final grade in a course in which the present authors were not involved. Five of the 25% was determined by anonymous peer-evaluations of a student's teamplaying behaviour.

The rationale for using project performance as an assessment of maximum performance is three-fold. First, consistent with the Sackett et al.'s (1988) requirement for awareness of one's evaluation throughout the evaluation period, the course-instructor informed the students of the on-going peer assessment of teamplaying performance throughout the 5-day period. Second, the instructor explained to the students the necessity to perform at one's best, and he reminded them that their teamplaying performance during the project would influence their course grade. Course grades are crucial in first year MBA program to attain a relevant summer job. Hence the students were under high pressure to perform at their best. These factors satisfy Sackett et al.'s (1988) second requirement for maximum performance. Third, the project lasted only one week. Relative to the assessment of typical performance throughout the entire semester, the students could focus their attention on this project.

Typical performance was defined as the arithmetic mean of scores given to a student by peers ( $n = 4-5$ ) within the person's study-group. Performance was not assessed until the end of students' first term in the MBA program, but before participants had received their grades. This measure satisfies Sackett et al.'s (1988) definition of typical performance in that: (a) the assessment represented students' mean performance over an extended time-period, namely four months as opposed to 5-days; (b) students were not aware during this time-period that they would be evaluated by their peers at the end of the term; and (c) the students were not given any instructions by the faculty during that period to maximise their effort.

This operationalization of typical versus maximum performance allowed for comparable assessments. Both assessments were conducted in temporal proximity to each other, namely, before the end of the participants' first term in the MBA program. In both cases, peers evaluated participants on exactly the same behaviours using the same appraisal scale. To minimise the possibility that the assessments of maximum performance could influence the assessments of typical performance taken a few days later, no two appraisers from the same study-group (typical performance) were members of the same-project group (maximum performance).

### Manipulation checks

A serious limitation of most studies of typical and maximum performance

(e.g., Barnes & Morgeson, 2007; ForsterLee, 2007; Klehe & Anderson, 2007b; Mangos et al., 2007; Marcus et al., 2007; Ones & Viswesvaran, 2007; Ployhart et al., 2001; Sackett et al., 1988; Smith-Jentsch, 2007; Witt & Spitzmüller, 2007) is the lack of manipulation checks. Immediately after the respective assessments of maximum and typical performance in the present study, the participants indicated on a scale from 0 (*almost never*) to 4 (*almost always*): (a) how often they had been aware that they were being evaluated; (b) how often they had done their very best to contribute to the group's performance; and (c) how focused they had been on their group's work.

In addition, we examined whether the behaviours assessed by the BOS were of comparable relevance in the typical and maximum conditions in order to ensure comparable dimensionality of performance. We also examined the ease with which the behaviours could be observed in the typical and maximum performance situations to ensure comparable reliability of performance observation (Cascio, 1998). Thus MBA-students, enrolled in the same business school, who were not involved in the validation study, evaluated the BOS-items on relevance and ease of observation in either their respective study groups ( $n = 51$ ), or during their one-week project group ( $n = 53$ ). Relevance and observability were assessed on a 5-point Likert-type scale, ranging from 0 (*not at all*) to 4 (*very much*). The power of detecting a medium effect size on either rating, ease of observation or relevance of the 13 BOS-items in the typical and maximum performance situation is .80 (Cohen, 1988; 1992).

#### Tests of maximum performance – task-related knowledge

Participants' task related abilities (task-related knowledge and skills) were assessed during the participants' second week in the MBA program. Teamplaying related job knowledge was assessed with a 35 item multiple-choice test developed by Stevens and Campion (1999) to assess the knowledge required to be a teamplayer. In a study involving workers in a pulp-mill and a cardboard box plant ( $n = 142$ ), Stevens and Campion found that the internal consistency of the test was high ( $\alpha = .80$ ), and that the test correlated significantly with supervisory ratings of both teamwork and performance ( $r = .32, p < .01, r = .37, p < .01$ , respectively).

#### Tests of maximum performance – skills

Task related skills were assessed in the form of a practical intelligence test developed by Sue-Chan, Latham, and Evans (1995). It consists of 6 items that describe specific study-related situations, each accompanied by 7 or 8 strategies for dealing with the situation. Participants were asked to rate each strategy from 1 (*extremely useless*) to 7 (*extremely useful*). Each participant's score was calculated as the sum of the absolute difference between the rating

of each response strategy, and the mean rating of the response strategy provided by a group of SMEs (Sue-Chan et al., 1995). The SMEs in the present study were 26 advanced MBA-students from the same MBA program as the participants, but from an earlier cohort. For this purpose, the school's MBA-student association recruited successful students who were near the end of their 2-year MBA program. We chose nearly graduated students as SMEs, since successful students at the end of the MBA program were likely to have excellent insight as to the behaviours that were relevant, desirable, and manageable in an MBA program.

#### Predicting tests of maximum performance – general mental ability

General mental ability was assessed on the basis of the student's GMAT-score, which combines measures of verbal and quantitative ability (ETS, 2001).

#### Tests of typical performance – past decisions and future intentions

We chose a set of structured past- (BDI) and future-oriented (SI) interview-questions as tests of typical performance. The reason for classifying structured interviews as tests of typical performance lies in Cronbach's (1960) notion that tests of typical performance attempt to address what a person would do rather than what that person could do on the job. In line with this definition, both forms of structured interviews, the BDI and SI typically tap respondents' motivation. The underlying assumption of the PBDI is that the answers "reveal specific choices applicants have made in the past, and the circumstances surrounding those choices" (Janz, 1989, p. 159). Janz (1989, pp. 159-160) argued that "the more long-standing the applicant's behaviour pattern in the past, the more likely it will predict behaviour in the future". Similarly, Latham (1989) argued that SI questions measure intentions that predict behaviour. In order to ensure that respondents' answers truly reflect their intentions rather than ability to give socially desirable responses, SI questions contain a dilemma for which the correct answer is not evident. Thus, when confronted with situational dilemmas, interviewees are 'forced' to state their behavioural intentions (Latham & Skarlicki, 1995; Maurer, Sue-Chan, & Latham, 1999). Intentions, a core variable in social cognitive theory (Bandura, 1986), are "a representation of a future course of action to be performed ... a proactive commitment to bringing them (future actions) about" (Bandura, 2001, p. 5). They are generally viewed as the direct motivational instigator of behaviour (Fishbein & Ajzen, 1975; Lewin, 1951; Locke & Latham, 1990).

Four 2nd-year MBA students served as SMEs in collecting 24 critical incidents on teamplaying behaviour during the MBA program. On the basis of these incidents, they developed 16 parallel interview questions that

focused either on choices candidates had made in the past ("What did you do?") or on choices they believed they would make in the future when facing different work-related situations ("What would you do?").

SMEs also developed parallel evaluation guides with behavioural benchmarks for outstanding (5), acceptable (3), and unacceptable (1) answers. Following a pilot test with 31 management students who were not involved in the current study, the SMEs, together with eight doctoral students in human resource management, revised the scoring guide and eliminated questions that elicited unclear answers or answers with little or no variance in responses. This resulted in nine comparable past- and future-oriented questions.

Interviewees were recruited for the interviews before the start of their studies via a letter sent to them with their welcoming package to the MBA program. With the exception of the GMAT, no selection decisions were made on the basis of the two interviews or any other predictor test involved in the current study. Nevertheless, participants' motivation during the two interviews was high, as was obvious from the formal manner with which interviewees approached the interview-situation and the frequent request for feedback on their behavioural interviews. At the beginning of both interviews, interviewees learned that they would be asked questions about difficult situations that they were likely to encounter during the MBA program. Each participant was then interviewed by the first author and a doctoral student who had been trained as an interviewer. The latter individual was blind to the purpose of the study. This individual read the questions to the participant, and both interviewers independently recorded and scored the responses on the basis of a scoring guide. Following each interview, the two interviewers discussed the scoring of each response on which they differed by more than 1 point on the 5-point scale.

## Results

### *Manipulation checks and descriptives*

Independent two tailed  $t$  tests on the relevance ( $M = 3.09$ ,  $SD = .63$  vs.  $M = 3.17$ ,  $SD = .41$ ;  $t = .70$ ,  $ns$ ) as well as the ease of observation-ratings ( $M = 2.71$ ,  $SD = .58$  vs.  $M = 2.83$ ,  $SD = .58$ ;  $t = 1.12$ ,  $ns$ ) revealed no differences between the typical and maximum performance conditions on the 13 BOS items. The relevance ( $t = 4.10$ – $18.09$ ,  $p > .01$ ) and ease of observation ( $t = 2.39$ – $16.84$ ,  $p > .05$ ) of all items were always rated above 2, the graphic midpoint of the Likert-type scale.

Immediately after the assessments of maximum and typical performance,

Table 1  
*Correlation Matrix*

	<i>n</i>	<i>M</i>	<i>S.D.</i>	Age	Gender	GMA	PI	KSA	Future intent.	Past decision	Typ. Perf.	Max. Perf.
Control variables												
Age	167	28.18	3.48	-	.11	-.05	-.16	-.05	-.08	-.21	-.27**	-.27**
Gender <sup>a</sup>	167	0.31	.46	-.03	-	-.34**	.05	.08	.13	-.10	-.02	.05
Ability												
General mental ability	167	665.45	48.37	-.10	-.28**	-	.05	.10	.04	.13	.09	.01
Practical												
Intelligence (PI)	120	63.35	12.95	-.15	.05	.07	(.75)	.42**	.42**	.24	.23	.31**
Task knowledge	157	23.38	4.47	-.21**	.03	.16*	.39**	(.88)	.40**	.40**	.27**	.38**
Motivation												
Future intentions	79	3.00	.47	-.15	.07	.07	.45**	.42**	(.50)	.48**	.40**	.37**
Past decisions	79	2.68	.59	-.18	.01	.14	.26*	.36**	.48**	(.71)	.34**	.19
Performance												
Typical performance	162	3.18	.54	-.24**	.12	.20**	.24*	.32**	.41**	.34**	(.96)	.29*
Maximum performance	167	3.37	.45	-.26**	.08	.12	.38**	.40**	.25*	.11	.34**	(.96)

*Note.* \*  $p < .05$ ; \*\*  $p < .01$ ; Internal consistencies are reported in the diagonal. Correlations below the diagonal represent results after pairwise deletion of missing data, correlations above the diagonal represent results after listwise deletion of missing data.  
a 0 = male, 1 = female, correlations are point biserial, since the variable is dichotomous.

participants had indicated the level of typical versus maximum performance on three items about their perceptions of evaluation, their focus and effort to do one's best. A paired  $t$  test showed that on average, the two performance measures truly assessed typical ( $M = 2.90$ ,  $SD = .56$ ) versus maximum performance ( $M = 3.16$ ,  $SD = .53$ ;  $t = 3.47$ ,  $p < .01$ ). In sum, these results indicate that participants perceived the maximum condition as more "maximum" than the typical performance condition, and that the assessments of typical and maximum performance did not differ in regard to relevance or observability. Thus, the manipulation of typical versus maximum performance was successful. Table 1 presents the means, standard-deviations, internal consistencies and intercorrelations among the variables in this study.

### *Hypothesis 1*

Hypothesis 1 was tested via a comparison of measurement models with AMOS 5.0 (Arbuckle, 2003). The proposed measurement model suggested that the two tests of typical performance, namely past decisions and future intentions, loaded onto the latent 'task-related motivation' factor, and that the two tests of maximum performance, namely the tests of task related knowledge and skills (teamplaying KSAs and practical intelligence) loaded onto the latent 'task-related ability' factor, while allowing the two latent factors to correlate with one another. In addition, following earlier theory and empirical findings (e.g., Campbell, 1990; Campbell et al., 1996; Campbell et al., 1993; Hunter, 1983), we assumed that general mental ability would form a separate factor correlated with, but not loading onto the task-related ability factor.

In order to support Hypothesis 1, a 3-factorial measurement model (Model 1) should fit the data significantly better than a one-factorial model with all measures loading onto the same factor (Model 2). To test the models' fit, we used the overall model  $\chi^2$  measure and the  $\chi^2/df$  ratio, which should be below three (Byrne, 1994; 1998). An acceptable fit is further indicated by a goodness of fit index (GFI) and adjusted goodness of fit index (AGFI), as well an incremental fit index (IFI), a Tucker-Lewis index (TLI) and a comparative fit index (CFI) of at least .90, and by a root mean square error of approximation (RMSEA) of at most .08 (Byrne, 1998).

The proposed three-factor model (Model 1; see Table 2) showed an excellent fit to the data, a fit that was significantly better than a one-factor Model (Model 2;  $\Delta\chi^2_{(df=2)} = 8.74$ ;  $p < .01$ ). Based on the relatively high correlations between teamplaying KSAs, practical intelligence, future intentions, and past decisions, we also included an additional model comparison without the inclusion of general mental ability into the model. Here, a two-factor solution (teamplaying KSAs and practical intelligence loading onto one factor, and future intentions and past decisions loading onto a second factor; Model 3) produced

Table 2  
*Goodness-of-Fit Indices and Model Comparisons for the Measurement Models tested*

Model	$\chi^2$	df	p	$\chi^2/df$	GFI	AGFI	TLI	IFI	CFI	RMSEA	RMSEA upper 90%	Model comparison	$\Delta \chi^2$	df	P
1: 3 factor measurement model	3.01	3	.39	1.00	.99	.96	1.00	1.00	1.00	.00	.13				
2: 1 factor measurement model	11.75	5	.04	2.35	.97	.92	.87	.94	.94	.09	.16	Model 2 – Model 1	8.74	2	<.01
3: without GMA – 2 factor measurement model	1.30	1	.25	1.30	1.00	.96	.98	1.00	1.00	.04	.22				
4: without GMA – 1 factor measurement model	9.30	2	.01	4.65	.97	.86	.79	.93	.93	.15	.25	Model 4 – Model 3	8.00	1	<.01
5: proposed model	11.62	12	.48	.97	.98	.95	1.00	1.00	1.00	.00	.08				
6: full model	8.89	11	.63	.81	.98	.96	1.00	1.00	1.00	.00	.07	Model 5 – Model 6	2.73	1	.10
7: direct model	43.51	12	.00	3.63	.93	.85	.72	.85	.84	.13	.17				

*Note.* N = 167. GFI = Goodness of Fit Index; AGFI = Adjusted Goodness of Fit Index; TLI = Tucker-Lewis Index; IFI = Incremental Fit Index; CFI = Comparative Fit Index; RMSEA = Root Mean Square Error of Approximation.



a significantly better fit to the data than a one factor solution (Model 4;  $\Delta X^2_{(df=1)} = 8.00$ ;  $p < .01$ ). Again, this suggests that task-related abilities are a separate factor from task-related motivation. In sum, Hypothesis 1 was thus accepted.

*Hypotheses 2 to 4*

These three hypotheses were tested via structural equation modeling using AMOS 5.0 (Arbuckle, 2003). Figure 2 presents the results of the proposed model (Model 5). This model shows an admissible solution with an excellent fit to the data ( $X^2 = 11.62$ ,  $df = 12$ , n.s.,  $X^2/df = .97$ ; CFI = IFI = TLI = 1.00; GFI = .98; AGFI = .95; RMSEA = .00;  $RMSEA_{high90\%} = .08$ ). Consistent with past theory and empirical findings (Campbell, 1990; Campbell et al., 1996; Campbell et al., 1993; Hunter, 1983; Motowidlo et al., 1997), participants' general mental ability predicted participants' task-related ability ( $\gamma = .21$ ,  $p = .02$ ). Consistent with Hypothesis 2, participants' task-related ability predicted their motivation ( $\gamma = .62$ ,  $p < .01$ ). This suggests that the decisions or intentions formed in a particular work-related situation are affected by a person's knowledge, skills, and abilities. Consistent with Hypotheses 3 and 4, both ability ( $\gamma = .29$ ,  $p < .05$ ) and motivation ( $\gamma = .33$ ,  $p < .01$ ) predicted typical performance ( $R^2 = .32$ ), whereas only ability ( $\gamma = .65$ ,  $p < .01$ ) predicted maximum performance ( $R^2 = .42$ ). These data support the validity of the proposed model.

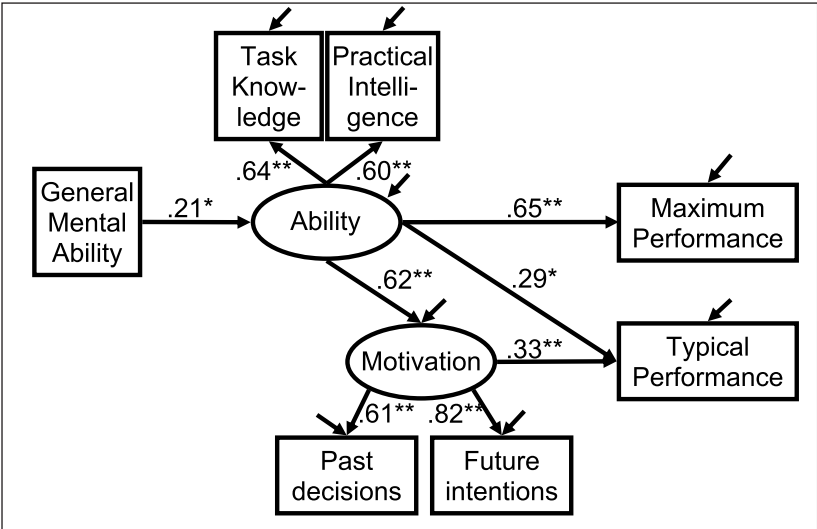


Figure 2  
*Results of the proposed model linking the predictors ability and motivation with the criteria of maximum and typical performance*

Additionally, we tested for indirect effects of ability on the possible link between motivation and maximum performance, and for the possible partial mediating effect of motivation on the link between ability and typical performance. For this purpose, we ran two alternative models. The first of these two models represents a full model in which both task-related ability and motivation predict both maximum and typical performance (Model 6). The results indicate that the additional path from motivation to maximum performance did not reach significance. The second model represented a direct model where the direct link between task-related ability and motivation was deleted. This allowed for an assessment of direct effects between motivation and task-related abilities on the predictor, and typical and maximum performance on the criterion side (Model 7). The respective model showed a considerable decrement in fit. Moreover, the direct link from motivation to maximum performance again failed to reach significance. This suggests there is no link between motivation and maximum performance. The direct link between ability and typical performance, however, was highly significant. Nevertheless, the path-weight ( $\gamma = .34, p < .01$ ) barely surpassed the path-weight found in the proposed model (Model 5). A Sobel-test indicated that the mediating effect was not significant ( $t = 1.41; p = .16$ ). This suggests that the impact of task-related ability on typical performance is not a function of participants' motivation.

### *Hypothesis 5*

A Pitman-Morgan test (Morgan, 1939; Pitman, 1939) assessed the effect of typical versus maximum performance conditions on the variance of performance. While the variance in the typical performance condition ( $\sigma^2 = .23$ ) was somewhat larger than the variance in the maximum performance condition ( $\sigma^2 = .20$ ), the 95%-confidence interval around the ratio between standard-deviations ranged from .96 to 1.45, thus including the 1. This suggests that the difference in variance was not significant. Hypothesis 5a was thus rejected. In order to test Hypothesis 5b, a paired t-test compared the ratings participants received on typical versus maximum performance. In support of Hypothesis 5b, participants' maximum performance ( $M = 3.37, SD = .45$ ) significantly surpassed their typical performance ( $M = 3.20, SD = .47; t_{(df=166)} = 4.28; p < .01$ ).

## Discussion

The current study addressed the prediction of typical versus maximum performance (Sackett et al., 1988) with task-related tests of typical and maximum performance (Cronbach, 1960). The results confirmed that predictors of maximum performance, that is, measures of task-related ability, loaded onto a different factor than tests of typical performance, that is measures of task-related motivation. This occurred even though these two factors correlated moderately with one another. Further, typical teamplaying performance was predicted by both participants' task related ability and motivation. Maximum teamplaying performance, on the other hand, was only predicted by task-related ability. Thus, the current study provides the first coherent link between the literature on tests of typical versus maximum performance on the predictor side (Cronbach, 1960) and the literature on typical versus maximum performance on the criterion side (Sackett, 2007; Sackett et al., 1988).

Unlike previous studies (e.g., Ployhart et al., 2001), the predictors used in the current study had been explicitly developed to predict the criterion targeted in the current study, namely teamplaying. Earlier field studies on typical versus maximum performance often employed relatively broad predictors of performance, such as general mental ability one or more of the Big-Five (e.g., Ployhart et al., 2001; Witt & Spitzmüller, 2007). Other studies have assessed motivation and ability at different levels of task related specificity (e.g., Ones & Viswesvaran, 2007). This made a comparison of the two difficult (see Marcus et al., 2007 for a more refined approach).

Tailoring the predictors to the criteria likely contributed to the relatively high prediction of both typical versus maximum performance, even though the manipulation of typical versus maximum performance was considerably weaker in the current study than in earlier research. As Guion (1991) noted, researchers often run extensive validation studies, and organisations often make huge financial investments in the selection of employees without knowing which of these two aspects of performance they are predicting, or even trying to predict (Klehe & Anderson, 2005). The present results suggest that for jobs requiring typical performance, structured interviews may be the predictor of choice. For predicting maximum performance, as is relevant for jobs such as fire-fighters or other emergency-related personnel, measures of task-related knowledge are useful.

While the manipulation checks ruled out alternative explanations for the results, the present results do not support all the basic assumptions underlying the distinction between typical and maximum performance. The participants' maximum performance did surpass participants' typical performance as had been hypothesised. Nevertheless, the variances between these two measures of performance did not significantly differ from one another. A

significant difference was hypothesised based on Sackett et al.'s (1988; see also DuBois et al., 1993) assertion that maximum performance situations *constrain* a performer's motivation to be high. The results of the current study at best point in the proposed direction, but they are not significant. Reanalyses of earlier studies (e.g., Klehe et al., 2007; Witt & Spitzmüller, 2007) suggest that this null finding is not atypical and that results might sometimes even point in the direction opposite to the one proposed by Sackett et al. (e.g., Barnes & Morgeson, 2007; Klehe & Anderson, 2007b; Lim & Ployhart, 2004; Ployhart et al., 2001).

The question remains as to when maximum performance reduces variance in performance. Klehe and Anderson (2007b) found that people worked harder under maximum performance conditions; they worked smarter under typical performance conditions. In another study, Klehe, Anderson, and Hoefnagels (2007) found that maximum performance conditions triggered anxiety which in turn inhibited performance. Individual differences regarding susceptibility to anxiety may increase the variance in a maximum performance situation (McCarthy & Goffin, 2004). Future research on this issue is necessary for at least two reasons. First, a realistic set of factors that explain performance under typical and maximum performance conditions is needed. Second, the emotions and cognitions that affect people under maximum performance conditions need to be understood in order to minimise counterproductive factors that affect a person's well-being (Barling, Kelloway, & Frone, 2005).

### *Limitations and directions for future research*

A limitation of this study is the setting. Data collected in a non-profit educational setting may not yield results that generalise to private sector for-profit organisations. Nevertheless, the present study took place in a high-pressure learning-environment. The performance outcomes (i.e., the grades) were relevant for the individual and the group to which the individual belongs. This setting raised participants' personal stakes in attaining high typical performance. This fact likely reduced the difference between typical and maximum performance. This in turn should have reduced the likelihood of finding significant differences. That this was not the case suggests the robustness of the present findings.

A second limitation lies in the assessment of motivation. Motivation is usually defined in terms of the choice, level, and persistence of effort (Campbell, 1990; Sackett, 2007). Effort and persistence were not directly measured in this study. The motivational predictors used in the current study included a behaviour description interview that focused on choices made in the past (Janz, 1989) and a situational interview that focused on future inten-

tions (Latham, 1989). Both methods only assess the choices people make when deciding between different courses of possible actions, and thus on the direction of effort. However, Klehe and Anderson (2007b) found that choice and level of effort are highly correlated. To predict persistence, personality tests of conscientiousness may prove useful in future research (Costa & McCrae, 1992).

The tests of typical and maximum performance in this study were assessed in a predictive validation study. Yet with the exception of the GMAT, people were not selected on the basis of these predictors. Administering the tests in a non-selection context may have reduced participants' motivation during the test to do as well as possible. Yet, as noted earlier, the researchers' observations of the interviewees suggested that they took the interviews seriously.

### References

- Ackerman, P.L., & Humphreys, L.G. (1990). Individual differences in industrial and organizational psychology. In M.D. Dunnette & L.M. Hough (Eds.), *Handbook of Industrial and Organizational Psychology* (Vol. 1 (2nd ed.), pp. 223-282). Palo Alto, CA: Consulting Psychologists Press, Inc.
- Allred, B.B., Snow, C.C., & Miles, R.E. (1996). Characteristics of managerial careers in the 21st century. *Academy of Management Executive*, 10, 17-27.
- Arbuckle, J.L. (2003). *AMOS* (5 ed.). Chicago, IL: Smallwaters Corporation.
- Arvey, R.D., & Murphy, K.R. (1998). Performance evaluation in work settings. *Annual Review of Psychology*, 49, 141-168.
- Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Upper Saddle River, NJ, US: Prentice-Hall, Inc.
- Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: Freeman.
- Bandura, A. (2001). Social cognitive theory: An agentic perspective. *Annual Review of Psychology*, 52, 1-26.
- Barling, J., Kelloway, E.K., & Frone, M.R. (Eds.) (2005). *Handbook of work stress*. Thousand Oaks, CA: Sage.
- Barnes, C.M., & Morgeson, F.P. (2007). Typical performance, maximal performance, and performance variability: Expanding our understanding of how organizations value performance. *Human Performance*, 20, 259-274.
- Borman, W.C. (1991). Job behavior, performance, and effectiveness. In M.D. Dunnette & L.M. Hough (Eds.), *Handbook of industrial and organizational psychology* (Vol. 2 (2nd ed.), pp. 271-326). Palo Alto, CA: Consulting Psychologists Press, Inc.
- Boudreau, J.W. (1991). Utility analysis for decisions in human resource management. In M.D. Dunnette & L.M. Hough (Eds.), *Handbook of industrial and organizational psychology* (Vol. 2 (2nd ed.), pp. 621-745). Palo Alto, CA: Consulting Psychologists Press, Inc.
- Brodbeck, F.C., Frese, M., Akerblom, S., Audia, G., Bakacsi, G., Bendova, H., et al.

- (2000). Cultural variation of employeeship prototypes across 22 European countries. *Journal of Occupational & Organizational Psychology*, 73, 1-29.
- Byrne, B.M. (1994). *Structural equation modeling with EQS and EQS/WINDOWS. Basic concepts, applications, and programming*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Byrne, B.M. (1998). *Structural equation modeling with LISREL, PRELIS, and SIMPLIS: basic concepts, applications, and programming*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Campbell, J.P. (1990). Modeling the performance prediction problem in industrial and organizational psychology. In M.D. Dunnette & L.M. Hough (Eds.), *Handbook of industrial and organizational psychology* (Vol. 1 (2nd ed.), pp. 687-732). Palo Alto, CA: Consulting Psychologists Press Inc.
- Campbell, J.P., Gasser, M.B., & Oswald, F.L. (1996). The substantive nature of job performance variability. In K.R. Murphy (Ed.), *Individual differences and behavior in organizations* (pp. 258-299). San Francisco, CA: Jossey-Bass.
- Campbell, J.P., McCloy, R.A., Oppler, S.H., & Sager, C.E. (1993). A theory of performance. In N. Schmitt & W.C. Borman (Eds.), *Personnel selection in organizations* (pp. 35-70). San Francisco, CA: Jossey Bass.
- Cascio, W.F. (1998). *Applied psychology in human resource management* (5th ed.). Upper Saddle River, NJ: Prentice-Hall.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ, England: Lawrence Erlbaum Associates.
- Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112, 155-159.
- Costa, P.T., & McCrae, R.R. (1992). *Revised NEO Personality Inventory (NEO-PI-R) and NEO Five-Factor Inventory (NEO-FFI) professional manual*. Odessa, FL: Psychological Assessment Resources.
- Cronbach, L.J. (1949). *Essentials of psychological testing*. New York: Harper.
- Cronbach, L.J. (1960). *Essentials of psychological testing* (2nd ed.). Oxford, England: Harper.
- Dewberry, C. (2001). Performance disparities between whites and ethnic minorities: Real differences or assessment bias? *Journal of Occupational and Organizational Psychology*, 74, 659-673.
- DuBois, C.L.Z., Sackett, P.R., Zedeck, S., & Fogli, L. (1993). Further exploration of typical and maximum performance criteria: Definitional issues, prediction, and white-black differences. *Journal of Applied Psychology*, 78, 205-211.
- ETS. (2001). *GMAT Examinee Score Interpretation Guide*. <http://www.gmac.com/publications/rad0E9F4.pdf> [out of order].
- Fishbein, M., & Ajzen, I. (1975). *Belief, attitude, intention and behavior: An introduction to theory and research*. Reading, MA: Addison-Wesley.
- ForsterLee, R. (2007). Personality, interest, and motivational determinants of maximal and typical performances on a verbal knowledge task. *Human Performance*, 20, 287-292.
- Guion, R.M. (1991). Personnel assessment, selection, and placement. In M.D. Dunnette & L.M. Hough (Eds.), *Handbook of industrial and organizational psychology* (Vol. 2 (2nd ed.), pp. 327-397). Palo Alto, CA: Consulting Psychologists Press, Inc.
- Guion, R.M. (1998). *Assessment, measurement, and prediction for personnel deci-*

- sions. Mahwah, NJ, US: Lawrence Erlbaum Associates.
- Hunter, J.E. (1983). A causal analysis of cognitive ability, job knowledge, job performance and supervisory ratings. In F.J. Landy, S. Zedeck, & J. Cleveland (Eds.), *Performance measurement and theory* (pp. 257-266). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- Janz, T. (1989). The patterned behavior description interview: The best prophet of the future is the past. In R.W. Eder & G.R. Ferris (Eds.), *The employment interview: Theory, research, and practice* (pp. 158-168). Thousand Oaks, CA: Sage.
- Kanfer, R., & Ackerman, P.L. (1989a). Dynamics of skill acquisition: Building a bridge between intelligence and motivation. In R.J. Sternberg (Ed.), *Advances in the psychology of human intelligence* (Vol. 5, pp. 83-134). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Kanfer, R., & Ackerman, P.L. (1989b). Motivation and cognitive abilities: An integrative/aptitude/treatment interaction approach to skill acquisition. *Journal of Applied Psychology*, 74, 657-690.
- Kanfer, R., Ackerman, P.L., Murtha, T.C., Dugale, B., & Nelson, L. (1994). Goal setting, conditions of practice, and task performance: A resource allocation perspective. *Journal of Applied Psychology*, 79, 826-835.
- Kirk, A.K., & Brown, D.F. (2003). Latent constructs of proximal and distal motivation predicting performance under maximum test conditions. *Journal of Applied Psychology*, 88, 40-49.
- Klehe, U.-C., & Anderson, N. (2005). The prediction of typical and maximum performance. In A. Evers, N. Anderson, & O. Smit-Voskuijl (Eds.), *Handbook of personnel selection* (pp. 331-353). Oxford, UK: Blackwell.
- Klehe, U.-C., & Anderson, N. (2007a). The moderating influence of personality and culture on social loafing in typical versus maximum performance situations. *International Journal of Selection and Assessment*, 15, 250-262.
- Klehe, U.-C., & Anderson, N. (2007b). Working hard and working smart: Motivation and ability during typical and maximum performance. *Journal of Applied Psychology*, 92, 978-992.
- Klehe, U.-C., Anderson, N., & Hoefnagels, E.A. (2007). Social facilitation and inhibition during maximum versus typical performance situations. *Human Performance*, 20, 223-239.
- Klehe, U.-C., Anderson, N., & Viswesvaran, C. (2007). More than peaks and valleys: Introduction to the special issue on typical and maximum performance. *Human Performance*, 20, 173-178.
- Klehe, U.-C., & Latham, G.P. (2006). What would you do – really or ideally? Constructs underlying the behavior description interview and the situational interview in predicting typical versus maximum performance. *Human Performance*, 19, 357-382.
- Latham, G.P. (1989). The reliability, validity, and practicality of the situational interview. In R.W. Eder & G.R. Ferris (Eds.), *The employment interview: Theory, research, and practice* (pp. 169-182). Thousand Oaks, CA: Sage.
- Latham, G.P., & Skarlicki, D.P. (1995). Criterion-related validity of the situational and patterned behavior description interviews with organizational citizenship behavior. *Human Performance*, 8, 67-80.
- Latham, G.P., & Wexley, K.N. (1977). Behavioral observation scales for performance



- appraisal purposes. *Personnel Psychology*, 30, 255-268.
- Latham, G.P., & Wexley, K.N. (1994). *Increasing productivity through performance appraisal* (2nd ed.). Reading, MA: Wesley Publishing Company.
- Lewin, K. (1951). *Field theory in social science: selected theoretical papers* (Edited by Dorwin Cartwright). Oxford, England: Harpers.
- Lim, B.C., & Ployhart, R.E. (2004). Transformational leadership: Relations to the Five-Factor model and team performance in typical and maximum contexts. *Journal of Applied Psychology*, 89, 610-621.
- Locke, E.A., & Latham, G.P. (1990). *A theory of goal setting & task performance*. Upper Saddle River, NJ, US: Prentice-Hall.
- Locke, E.A., Mento, A.J., & Katcher, B.L. (1978). The interaction of ability and motivation in performance: An exploration of the meaning of moderators. *Personnel Psychology*, 31, 269-280.
- Love, K.G. (1981). Comparison of peer assessment methods: Reliability, validity, friendship bias, and user reaction. *Journal of Applied Psychology*, 66, 451-457.
- Maier, N.R.F. (1955). *Psychology in industry* (2nd ed.). Oxford, England: Houghton Mifflin.
- Mangos, P.M., Steele-Johnson, D., LaHuis, D., & White, E.D. (2007). A multiple-task measurement framework for assessing maximum-typical performance. *Human Performance*, 20, 241-258.
- Marcus, B., Goffin, R.D., Johnston, N.G., & Rothstein, M.G. (2007). Personality and cognitive ability as predictors of typical and maximum managerial performance. *Human Performance*, 20, 275-285.
- Maurer, S.D., Sue-Chan, C., & Latham, G.P. (1999). The situational interview. In R.W. Eder & M.M. Harris (Eds.), *The employment interview handbook* (pp. 159-177). Thousand Oaks, CA: Sage.
- McCarthy, J., & Goffin, R. (2004). Measuring job interview anxiety: Beyond weak knees and sweaty palms. *Personnel Psychology*, 57, 607-637.
- Mesmer-Magnus, J., & Viswesvaran, C. (2007). Inducing maximal versus typical learning through the provision of a pretraining goal orientation. *Human Performance*, 20, 205-222.
- Morgan, W.A. (1939). A test for the significance of the difference between the two variances in a sample from a normal bivariate population. *Biometrika*, 31, 13-19.
- Motowidlo, S.J., Borman, W.C., & Schmit, M.J. (1997). A theory of individual differences in task and contextual performance. *Human Performance*, 10, 71-83.
- Mumford, M.D. (1983). Social comparison theory and the evaluation of peer evaluations: A review and some applied implications. *Personnel Psychology*, 36, 867-881.
- Ones, D.S., & Viswesvaran, C. (2007). A research note on the incremental validity of job knowledge and integrity tests for predicting maximal performance. *Human Performance*, 20, 293-303.
- Pitman, E.J.G. (1939). A note on normal correlation. *Biometrika*, 31, 9-12.
- Ployhart, R.E., Lim, B.C., & Chan, K.Y. (2001). Exploring relations between typical and maximum performance ratings and the five factor model of personality. *Personnel Psychology*, 54, 809-843.
- Prentice, D.A., & Miller, D.T. (1992). When small effects are impressive. *Psychological Bulletin*, 112, 160-164.
- Sackett, P.R. (2007). Revisiting the origins of the typical-maximum performance dis-



- tion. *Human Performance*, 20, 179-185.
- Sackett, P.R., Zedeck, S., & Fogli, L. (1988). Relations between measures of typical and maximum job performance. *Journal of Applied Psychology*, 73, 482-486.
- Smith-Jentsch, K.A. (2007). The impact of making targeted dimensions transparent on relations with typical performance predictors. *Human Performance*, 20, 187-203.
- Smith-Jentsch, K.A., Salas, E., & Brannick, M.T. (2001). To transfer or not to transfer? Investigating the combined effects of trainee characteristics, team leader support, and team climate. *Journal of Applied Psychology*, 86, 279-292.
- Stevens, M.J., & Campion, M.A. (1999). Staffing work teams: Development and validation of a selection test for teamwork settings. *Journal of Management*, 25, 207-228.
- Sue-Chan, C., & Latham, G.P. (2004). The relative effectiveness of external, peer, and self coaches. *Applied Psychology: An International Review*, 53, 260-278.
- Sue-Chan, C., Latham, G.P., & Evans, M.G. (1995). *The construct validity of the situational and patterned behavior description interviews: Cognitive ability, tacit knowledge, and self-efficacy as correlates*. Paper presented at the Annual meeting of the Canadian Psychological Association, Charlottetown, PEI.
- Thorsteinson, T.J., & Balzer, W.K. (1999). Effects of coworker information on perceptions and ratings of performance. *Journal of Organizational Behavior*, 20, 1157-1173.
- Tziner, A., Joanis, C., & Murphy, K.R. (2000). A comparison of three methods of performance appraisal with regard to goal properties, goal perception, and ratee satisfaction. *Group & Organization Management*, 25, 175-190.
- Tziner, A., & Kopelman, R. (1988). Effects of rating format on goal-setting dimensions: A field experiment. *Journal of Applied Psychology*, 73, 323-326.
- Tziner, A., Kopelman, R.E., & Livneh, N. (1993). Effects of performance appraisal format on perceived goal characteristics, appraisal process satisfaction, and changes in rated job performance: A field experiment. *Journal of Psychology*, 127, 281-291.
- Witt, L.A., & Spitzmüller, C. (2007). Person-situation predictors of maximum and typical performance. *Human Performance*, 20, 305-315.

Received June 27, 2008

Revision received October 18, 2008

Accepted October 19, 2008