

# Incremental Validity of Leaderless Group Discussion Ratings Over and Above General Mental Ability and Personality in Predicting Promotion

Xavier Borteyrou\*, Filip Lievens\*\*,  
Marilou Bruchon-Schweitzer\*\*\*,  
Anne Congard\*\*\*\* and Nicole Rascle\*\*\*\*\*

\*Université de Franche-Comte, France. xavier.borteyrou@univ-fcomte.fr

\*\*Ghent University, Belgium

\*\*\*Université de Bordeaux, France

\*\*\*\*Université d'Aix-Marseille, France

\*\*\*\*\*Université de Bordeaux, France

**Leaderless group discussions (LGDs) constitute one of the oldest assessment center exercises. In recent times, their added value has sometimes been questioned in light of trends to streamline assessment centers. The purpose of the present study is to examine the incremental validity of LGD ratings over cognitive ability scores and personality ratings for the prediction of extrinsic career success (i.e., promotion speed and number of promotions). We investigated this issue in the context of the promotion of French naval officers (N = 93) in an academy for high-level executive positions over a 10-year period. Results indicated that LGD ratings accounted for incremental variance in the prediction of promotion criterion measures, beyond cognitive ability and personality test scores. These results confirm that LGD ratings provide a unique contribution to the prediction of extrinsic career success in high-level executive positions.**

## 1. Introduction

The leaderless group discussion (LGD) was first developed in a military context (Bass, 1954). Prior to World War II, it originated in the German Army, particularly in the German Navy (Ansbacher, 1951). The technique was subsequently used in the British Navy and United States Army (Bass, 1954). The LDG was part of a larger program of officer assessment via traditional tests, situation tests, and other procedures. The LGD is based on the following principles: Several assesses are asked to engage in a discussion as a group for a given amount of time. No one is appointed as leader. Two or more assessors observe the interactions of the group and rate each member, usually via a checklist. The assessors do not participate in the discussion. The checklists typically include

items related to initiative, effectiveness in the presentation of ideas, and observable influence on the other members. As a result of quantifying the checklist ratings, a score is obtained, representing the individual member's status in that particular LGD group. Apart from these general principles, the design of LGDs also varies in terms of the number of assessors, the length of discussion time, and the directions given to the group (Gleason, 1957).

The LGD was and is still regarded as an important exercise in assessment centers alongside other exercises such as oral presentations, in-baskets, case studies, and role plays (Thornton & Mueller-Hanson, 2004). Eurich, Krause, Cigularov, and Thornton (2009) identified 43% of organizations reporting their use in United States. One reason is that the LGD enables to get a unique perspective on how candidates 'get ahead' (as indicated by leadership/

influencing behaviors) as well as 'get along,' (as indicated by prosocial/team oriented behaviors; Hogan & Shelton, 1998). Due to their unforeseen turns, LGDs might also demonstrate how candidates deal with unpredictable events and adapt to them.

However, various disadvantages of LGDs have also been noted. Some of these potential drawbacks are not new. For instance, already in 1961, Kaess, Witryol, and Nolan stated that 'it would appear almost impossible to meet standard conventions of objective testing in the LGD where the problem presented to the group is intentionally ambiguous, where the personal-social composition differs from group to group' (Kaess, Witryol, & Nolan, 1961, p. 345). Apart from the potential variability in topical discussions and group composition, other potential thorny issues deal with the cognitive demands placed on assessors when observing the fast interactions among multiple candidates. Despite these possible drawbacks, research documented that both the inter-rater reliabilities (Jones, 1981) and criterion-related validities of LGD ratings were satisfactory. For example, Robertson & Kandola (1982) found an uncorrected correlation coefficient of 0.34 between the performance obtained in an LGD and subsequent job performance. In addition, there is research showing that LGDs are less prone to coaching effects than other assessment center exercises (Kurecka, Austin, Johnson, & Mendoza, 1982; Lievens, 2002; Petty, 1974).

More recent concerns related to LGDs focus on their costs and logistics. It is then argued that nowadays it has become increasingly difficult to assemble at the same time a group of potentially geographically dispersed candidates. This has led to the development of online AC exercises (Tippins & Adler, 2011; Weekley, Hawkes, Guenole, & Ployhart, 2015) and to pressing questions whether more efficient procedures (e.g., personality inventories) might suffice for providing insights in how candidates get along (e.g., agreeableness) and get ahead (e.g., conscientiousness). Indeed, various studies have established links between personality and small group and leadership performance, respectively (Barry & Stewart, 1997; Bono & Judge, 2004; Waldman, Atwater, & Davidson, 2004).

In light of these utility questions inherent in LGDs, this study aims to examine the extent to which LGD ratings significantly add to the prediction above and beyond the Big Five personality trait and general mental ability. The present study is situated in the French Army. An assessment center was set up – for research and developmental purposes only – prior to French naval officers entering the Military Defense Academy. Once enlisted in the Military Defense Academy, all officers typically receive only high evaluations by their respective supervisors because all of them are considered to be elite officers. Accordingly, there exists virtually no variability in officers' job performance ratings. As the French army faces the problem of retaining 'high-potential' officers who are highly coveted by the defense industry (an industry that offers higher

salaries), promotions serve as an 'unofficial' approach for retaining the best of them. This promotion system in the French Navy corresponds to the notions underlying tournament theory (Lazear & Rosen, 1981). According to this theory, employees' extrinsic career success (e.g., promotion) is based on relative performances (people are ranked relative to one other) instead of on their performance and productivity on the job per se (which show little or no variance). In such contexts, it has been argued to use promotions as criterion measures for gauging the quality of the personnel (Kumazawa, 2010). Therefore, this study relied on promotion speed and the number of promotions as criterion measures (as indicators of extrinsic career success), thereby controlling for tenure.

## 2. Study background

Extrinsic career success 'refers to outcomes that are both instrumental rewards from the job or occupation' (Seibert & Kraimer, 2001, p. 2) and includes observable outcomes such as salary and promotions (Judge, Higgins, Thoresen, & Barrick, 1999; Ng, Eby, Sorensen, & Feldman, 2005; Seibert & Kraimer, 2001). Ng et al. (2005) conducted a meta-analysis in which they examined the relationship between various characteristics and objective and subjective career success. Stable individual differences such as the Big Five personality traits and cognitive ability were also scrutinized. Results showed that Extraversion and Conscientiousness positively predicted the number of promotions, whereas Emotional Stability and Agreeableness emerged as negative predictors. In the meta-analysis, there were not enough data points for estimating the validity of cognitive ability in relation to promotion. However, cognitive ability was a significant predictor of another objective career success criterion, namely salary. It was noteworthy that Judge et al. (1999) found exactly the same results. Apart from cognitive ability, the same four personality traits were enduring predictors of extrinsic career success, even over a time period of 50 years.

Besides personality and cognitive ability, assessment center ratings have also a long history of being used as instruments for predicting career progression (see the well-known Management Progress Study; Howard & Bray, 1988). Gaugler, Rosenthal, Thornton, and Bentson (1987) meta-analyzed the relationship between assessment center ratings and a composite of career advancement criteria (e.g., number of promotions, salary). They found a corrected validity coefficient of 0.36 between the overall assessment center rating and career advancement. Unfortunately, these results were not broken down by assessment center exercise. In addition, no incremental validity analysis was conducted. Therefore, the added value of LGD ratings over and above personality and cognitive ability in the prediction of extrinsic career success was not determined. More recent AC meta-analyses on

the level of assessment center dimensions do also not permit to ascertain the incremental variance in extrinsic career success explained by LGD ratings over and above personality and cognitive ability scores (Arthur, Day, Mcnelly, & Edens, 2003; Goffin, Rothstein, & Johnston, 1996; Kasten, Dayan, & Fox, 2002; Meriac, Hoffman, Woehr, & Fleisher, 2008).

At the backdrop of the above, this study investigates the incremental validity of LGD ratings over and above the Big Five personality traits and cognitive ability for predicting promotion criterion measures (promotion speed and number of promotions) as extrinsic career success indicators. In particular, we conduct two types of incremental validity analyses. The first traditional incremental validity strategy examines whether adding predictor methods explain extra variance in criteria of interest beyond that which is accounted for by other, less expensive predictor methods (Schmidt & Hunter, 1998). As the validity of cognitive ability test scores and self-report personality inventory ratings have been well-established (Hurtz & Donovan, 2000; McHenry, Hough, Toquam, Hanson, & Ashworth, 1990; Ree & Earles, 1991; Salgado, 1998; Schmidt, 2002) and given that these predictors are less expensive, we examine whether all dimensions measured by LGD provide incremental validity beyond these two predictors in explaining promotion criteria. We expect this to be the case because an LGD enables to observe actual behavior of candidates in a simulated situation. Such samples of behavior should add to the predictions made solely based on cognitive ability or self-reported personality ratings (Callinan & Robertson, 2000). This first strategy is especially insightful from a practical utility standpoint because the results allow streamlining a selection process. This leads to the following hypothesis:

*Hypothesis 1:* LGD dimension ratings will add significant incremental variance in predicting the speed of promotion and the number of promotions over and above cognitive ability scores and personality ratings.

However, this first incremental validity analytic strategy also comes with a conceptual drawback. That is, no meaningful conceptual conclusions can be drawn because the results might result from the LGD being a different selection procedure than a personality inventory or from LDG capturing constructs other than the Big Five personality traits. Therefore, we also conduct more construct-driven incremental validity analyses. In these analyses, the construct under investigation is held constant. Such analyses avoid the pitfall of confounding the content (constructs measured) with the methods (techniques used for measuring the specific content; see Arthur & Villado, 2008; Lievens, Harris, Van Keer, & Bisqueret, 2003) so that we are able to determine whether dimensions as measured by the LGD method provide incremental variance over

and above conceptually related constructs measured by a different method (e.g., a personality inventory).

These more construct-driven incremental validity analyses require that the dimensions assessed by this study LGD (i.e., influence, organizing and planning, and problem solving) are conceptually linked to personality traits. Lievens, Chasteen, Day, and Christiansen (2006) asked experts to make such conceptual linkages. Based on the expert ratings, we linked influence to Extraversion, organizing and planning to Conscientiousness, and problem solving to Openness. As cognitive ability is a general predictor, it was also always included in all the incremental validity analyses. In short, this leads to the following set of hypotheses:

*Hypothesis 2a:* The LGD dimension rating of influence will add significant incremental variance in predicting the speed of promotion and the number of promotions over and above cognitive ability scores and personality ratings on Extraversion.

*Hypothesis 2b:* The LGD dimension rating of organizing and planning will add significant incremental variance in predicting the speed of promotion and the number of promotions over and above cognitive ability scores and personality ratings on Conscientiousness.

*Hypothesis 2c:* The LGD dimension rating of problem solving will add significant incremental variance in predicting the speed of promotion and the number of promotions over and above cognitive ability scores and personality ratings on Openness.

### 3. Method

#### 3.1. Sample and procedure

The assessment center took place prior to entering the Military Defense Academy (Collège interarmées de Défense). This academy is the highest institution for training senior French army officers. The data originated from assessment centers organized over a 4-year time span (between 1998 and 2001) in which a sample of 125 French male Navy officers were assessed. The sample is homogeneous regarding education level and job experience. The officers were asked to participate in the assessment center for research (to make improvements to the selection process) and development purposes, thereby mitigating the possibility of criterion contamination.

During each assessment center session, three candidates participated over a one-day period in a variety of exercises including an in-basket, a role-play, and a LGD (always at the end of the day). In addition, all candidates were also administered a cognitive ability test battery and a personality inventory.

Assessors were both psychologists and military officers. They participated in approximately 10 hr of frame-

of-reference training (Lievens, 2001; Schleicher, Day, Mayes, & Riggio, 2002) prior to serving as assessors. The training focused on defining the dimensions, behavioral observation, and the appropriate scoring of effective and ineffective behaviors. To check the content of exercises and the behavioral checklists an experimental pilot assessment center was conducted in 1998. All the exercises were videotaped and constituted excellent material in the frame-of-reference training.

### 3.2. Measures

#### 3.2.1. Leaderless group exercise scores

The purpose of the LGD was to reach a solution regarding the allocation of budget lines that would suit all participants. The LGD was structured in that each of the three candidates was given an initial position and some supporting information to use in the presentation of his point of view. Each candidate was also provided with 45 min to read a scenario summarizing the problem and to write down issues and potential strategies. In the actual discussion, one candidate was responsible for the implementation of the personnel policy, another candidate was in charge of the budget, and the third candidate was responsible for military staff recruitment. As noted, assessors were both experienced psychologists and military officers. Each candidate was assessed by two or three assessors.

Behavioral checklists were based on a previous job analysis conducted in 1997 using the critical incident technique (Flanagan, 1954). Assessors recorded their observations using a 0-2 scale: 0 = behavior did not occur, 1 = behavior was exhibited by the candidate once and 2 = behavior occurred more than once. Each behavioral checklist was first filled separately by professional and nonprofessional raters. In a second step, psychologist and military officer(s) reached consensus on the ratings and completed together the behavioral checklist. In this study, candidate scores were calculated on the basis of the consensual checklist.

In previous research (Borteyrou, Bruchon-Schweitzer, Rasclé, & Collomb, 2006), exploratory factor analyses revealed that the behavioral checklist item ratings could best be represented by three underlying factors. The first factor included 21 items (e.g., 'propose adequate solutions,' 'decide at the appropriate time') describing 'the extent to which an individual gathers information, effectively analyzes data and information, generates viable options, ideas and solutions and generates and recognizes imaginative solutions' (Arthur et al., 2003, p. 135). This factor was thus labeled 'problem solving.' The second factor included 15 items (e.g., 'convince others that we are right,' 'question the inverse argument,' 'show that he's right'). Items loading this factor describe 'the extent to which an individual persuades others to do something or adopt a point a view to produce desired results and takes

action in which the dominant influence is one's own conviction rather than the influence of others' opinion' (p. 134). This second factor was labeled 'influencing others.' The third factor consisted of eight items concerning organizing and planning (e.g., 'organize teamwork,' 'search the consensus with others,' 'assign tasks'). Items loading on this factor describe 'the extent to which an individual systemically arranges his/her own work and resources as well as that of others for efficient task accomplishment' (p. 135). So, the emergence of these dimensions (i.e., 'problem solving,' 'influencing others,' and 'organizing and planning') corresponded well to several dimensions in Arthur et al.'s (2003) taxonomy of assessment center dimensions. The ratings on the items that loaded on these respective factors were summed to compute dimension scores on the LGD. Internal consistency reliability coefficients associated with the 'problem solving,' 'influencing others,' and 'organizing and planning' scores were 0.90, 0.86, and 0.75, respectively. This study used the checklist items that resulted from these analyses. We refer interested readers to Borteyrou et al. (2006) for more details about the psychometric characteristics of the LDG checklist scores.

#### 3.2.2. Personality measure

We used the French version of the NEO personality inventory that is based on the five-factor model of personality (Costa, McCrae, & Rolland, 1998). It measures Neuroticism, Extraversion, Openness to Experience, Agreeableness, and Conscientiousness. Each single factor/domain consists of six primary factors/facets, which are summed to form a total domain score. The inventory is composed of 240 self-descriptive statements. There are six eight-item facet scales designed to capture the qualities that are subsumed by these domains. Items are answered on a 5-point scale ranging from strongly agree (1) to strongly disagree (5), and scales are balanced to control for the effects of acquiescence. Internal consistency for the domain scales ranged from 0.76 to 0.93. In addition, 3- and 6-year retest coefficients of adults' scores ranged from 0.63 to 0.83 (Costa et al., 1998). In our study, an exploratory factor analysis showed that the facet scales typically had the highest loadings on their designated Big Five factors. As an exception, note that the Openness factor split into two factors (Griffin & Hesketh, 2004; Musesel, Winter, Gelléri, & Schuler, 2011), providing a six-factor solution. For consistency with prior research, we report only the global Openness score. The internal consistency reliability coefficients for the Big Five personality scores in Table 1 are satisfactory.

#### 3.2.3. Ability measure

We used the Culture Fair Intelligence Test (Cattell, 1974), which is a paper-and-pencil perceptual test that consists of abstract geometric forms. It can be considered one of the best available fluid intelligence measures (Colom &

Table 1. Descriptive statistics and correlations between study variables ( $N = 93$ )

	M	SD	1	2	3	4	5	6	7	8	9	10	11	12
<i>Control measure*</i>														
1. Tenure	19.10	3.04												
<i>Criterion measures</i>														
2. Promotion speed	3.54	0.76	-.11											
3. Number of promotions	1.66	0.74	-.06	.56**										
<i>Predictor measures</i>														
4. Cognitive ability	29.40	4.47	-.02	-.29**	.20*									
5. Neuroticism	68.28	16.09	.20*	-.23*	-.11	-.08	(.84)							
6. Extraversion	112.77	12.30	.23*	.09	.16	.01	-.04	(.81)						
7. Openness	107.67	12.36	.06	.23*	.17	.21*	-.05	.38**	(.89)					
8. Agreeableness	114.06	16.96	.24*	.08	.10	.02	.02	.19	.07	(.79)				
9. Conscientiousness	131.62	15.74	.15	.04	.20*	.11	-.27***	.28**	-.02	-.08	(.92)			
10. LGD problem solving	33.37	5.03	-.01	.22*	.25*	.18	.07	.13	.04	.19	-.04	(.90)		
11. LGD influencing others	18.24	3.50	-.05	.30**	.27**	.38**	-.08	.03	.01	-.04	.11	.16	(.86)	
12. LGD organizing and planning	12.22	2.61	.10	.38**	.37**	.40**	-.04	.14	.13	.04	.19	.37**	.34**	(.75)

Note: Cohort was not included in the matrix because is a nominal variable. \*\* $p < .01$  \* $p < .05$ . Reliability coefficients are in brackets.

García-López, 2003). It consists of four subtests that were selected in light of their correlation with Spearman's general mental capacity. In the first test (3 min, 13 items) candidates had to select from five choices the one that best completes a progressive series. In the second subtest (4 min, 14 items), candidates were required to identify from among five figures the one that is different from the others. The third subtest (3 min, 13 items) required candidates to complete a matrix of geometric figures by choosing from one of five figures presented. In the last subtest (2 min, 10 items), candidates' task was to select from five choices the one that duplicates a topological relationship among geometric forms that is the same as the one given.

### 3.2.4. Criterion measure

Given the lack of variance in job performance ratings, promotion speed and number of promotions in the Academy served as criterion measure. As noted above, these criterion measures can be conceptually placed under the umbrella term of extrinsic/objective career success (Ng et al., 2005). Promotion-related variables and other career advancements indicators have also often been used as criteria in prior assessment center research (Gaugler et al., 1987; Jansen & Stoop, 2001; Jansen & Vinkenburg, 2006) and other fields (Kirchmeyer, 2002; Kumazawa, 2010; Shipilov, Labianca, Kalnysh, & Kalnysh, 2014).

All participants were high-ranking French military officers at the time of the assessment center. Most of them were 'capitaine de corvette' (the first rank of commanding officer) or 'capitaine de frégate' (the second rank of commanding officer). In 2014, all participants were 'capitaine de vaisseau' (the third rank of commanding officer), or 'amiral' (the first rank of general officer). We counted the number of promotions. This criterion measure varied between 1 and 3. Regarding the speed of promotion cri-

terion, this was determined on the speed of promoting between the second and third ranks of commanding officer. This varied between 60 months (i.e., 5 years) and 114 months (i.e., more than 9 years). The mean promotion speed was 6.5 years (see Table 1). Given that higher values for promotion speed indicate less objective career success, we recoded this measure so that higher values are indeed indicative of promotion speed. As could be expected, there was a high positive correlation (0.56;  $p < .01$ ) between number of promotions and promotion speed.

The officers who had left the institution between 1998 and 2013 ( $n = 32$ ) were not taken into account in this study. So predictor-criterion analyses were conducted on a basis of 93 navy officers' data.

### 3.2.5. Control measures

We did our utmost best to control for various possible extraneous contaminating factors that might affect the speed of promotion and number of promotions. As noted above, the assessment center results were not divulged to the superiors of the officers. In addition, we controlled for cohort (as noted above, there were four cohorts) and tenure. Finally, note that during this period the officer promotion system was not affected by governmental budget cuts.

## 4. Results and discussion

### 4.1. Descriptive statistics

Table 1 presents the means, standard deviations, and correlations among this study's variables. In terms of the correlations between the different predictors, it was noteworthy that two<sup>1</sup> of the LGD dimensions were significantly associated with GMA. Conversely, the LGD

dimensions were not significantly related to the personality traits. These results suggest that this particular LGD taps especially into cognitive ability (see also Hoffman, Melchers, Blair, Kleinmann, & Ladd, 2011 p. 384).

The first rows of Table 1 display the bivariate correlations between the various predictors and criteria. Cognitive ability scores were significantly correlated with promotion speed ( $r = 0.29, p < .01$ ) and with number of promotions ( $r = 0.20, p < .05$ ), confirming the importance of ability measures in predicting extrinsic career success (Judge et al., 1999; Ng et al., 2005). Given that our sample is range restricted (all candidates are highly educated), this study's predictive validity coefficients are probably somewhat lower than in meta-analytic research with job performance as criterion (Schmidt & Hunter, 1998).

Concerning the predictive validity of the personality trait ratings, our results underline the importance of neuroticism ( $r = -.23, p < .05$ ) and openness ( $r = 0.23, p < .05$ ) in predicting promotion speed. The observed validity coefficient related to neuroticism is slightly higher than that generally found in previous meta-analyses (an average observed validity of .09; Salgado, 2003). A possible explanation is that neuroticism is a particularly relevant personality factor in a military setting because officers are regularly exposed to stressful and confusing situations. In this context, emotional stability can be envisaged as a resource that enables them to be more effective. Although Openness was the only personality trait that was not significantly related to promotion in the meta-analysis of Ng et al. (2005), it was a significant predictor in our study. Griffin and Hesketh (Griffin & Hesketh, 2004) posit that openness may be relevant only in certain situations or occupational groups, such as those characterized by novelty or complexity. Conscientiousness scores were associated with the number of promotion ( $r = 0.20, p < .05$ ) but there was no significant relationship between conscientiousness and promotion speed.

Regarding the predictive validity of the LGD dimension scores, our results demonstrate links between specific aspects of behavior assessed during this exercise and promotion speed and number of promotions. All three LGD ratings were significant predictors. More precisely, the observed correlation between the LDG rating on 'organizing and planning' and promotion speed is 0.38 ( $p < .01$ ), whereas the correlation between promotion speed and LGD ratings on 'influencing others' and 'problem solving' were 0.30 ( $p < .01$ ) and 0.22 ( $p < .05$ ), respectively. Correlations between LGD ratings and number of promotions were also significant.

#### 4.2. Tests of hypotheses

Hypothesis 1 posited that LGD dimension ratings would have incremental validity over personality and cognitive ability in predicting the two promotion criteria. A hier-

archical regression analysis was conducted. The two control variables were introduced in the first step to ensure that the year of assessment (cohort) and tenure did not affect the criterion measures. The cognitive ability test score was entered in the second step. In the third step, the Big Five personality trait ratings were entered. In the final step, three LGD dimension scores were entered to ascertain their additional variance over cognitive ability and personality factors.

Table 2 presents the results of the hierarchical regression analysis for the two promotion criteria. No effect of cohort or tenure was observed. In support of Hypothesis 1, LGD dimension scores accounted for a significant additional portion of the variance in promotion speed over cognitive ability and personality ( $\Delta R^2 = 0.10^{**}$ ). A similar result was found for number of promotions ( $\Delta R^2 = 0.13^{**}$ ). Note that the personality trait ratings did not provide significant incremental validity over cognitive ability scores for both promotion speed and number of promotions (respectively,  $\Delta R^2 = .07, ns$  and  $\Delta R^2 = .06, ns$ ). Although the added prediction is not significant, the size of the incremental variance of personality is consistent with previous findings. For instance, Meriac et al.'s (2008) meta-analysis obtained a similar added contribution of the five personality traits over cognitive ability ( $\Delta R^2 = .09$ ).

The second set of hypotheses (Hypotheses 2a, 2b, and 2c) proposed that a LGD rating on a specific dimension would have incremental validity for predicting the promotion criteria over and above conceptually related personality ratings (and cognitive ability). So, these hypotheses kept the construct under investigation constant. Table 3 summarizes the results of these incremental validity analyses. For example, in the regression related to the added

Table 2. Hierarchical regressions of promotion criterion measures on GMA, personality factors, and LGD dimensions

Predictors	Promotion speed		Number of promotions	
	Beta	R <sup>2</sup>	Beta	R <sup>2</sup>
<b>Step 1</b>				
Cohort	.05	.05	.03	.03
Tenure	-.01		-.14	
<b>Step 2</b>				
GMA	.06	.12	-.04	.06
<b>Step 3</b>				
Neuroticism	-.20		-.01	
Extraversion	-.01		.05	
Openness	.17		.12	
Agreeableness	.05		.22	
Conscientiousness	-.06		.18	
<b>Step 4</b>				
LGD problem solving	.07	.29	.13	.25
LGD influencing others	.15	.10*	.17	.13**
LGD organizing and planning	.26		.27*	

$N = 93$ . \* $p < .05$ . \*\* $p < .01$ . Cohort was coded into four values (1-2-3-4) in accordance with the year of assessment (1998, 1999, 2000, 2001).

Table 3. Construct-driven hierarchical regressions of promotion criterion measures on GMA, personality factors, and LGD dimensions

	Promotion speed			Number of promotions		
	Beta	R <sup>2</sup>	ΔR <sup>2</sup>	Beta	R <sup>2</sup>	ΔR <sup>2</sup>
<b>Step 1</b>		.05	.05		.03	.03
Cohort	.12			.07		
Tenure	-.16			-.11		
<b>Step 2</b>		.12	.07*		.06	.03
GMA	.20			.09		
<b>Step 3</b>		.12	.00		.08	.02
Extraversion	.08			.15		
<b>Step 4</b>		.17	.05*		.12	.04*
LGD Influencing others	.22			.22*		
<b>Step 1</b>		.05	.05		.03	.03
Cohort	.09			.05		
Tenure	-.16			-.13		
<b>Step 2</b>		.12	.07*		.06	.03
GMA	.11			.04		
<b>Step 3</b>		.12	.00		.09	.03
C	.14			.13		
<b>Step 4</b>		.21	.09**		.17	.08**
LGD Organizing and planning	.28*			.33*		
<b>Step 1</b>		.05	.05		.03	.03
Cohort	.12			.07		
Tenure	-.16			-.09		
<b>Step 2</b>		.12	.07*		.06	.03*
GMA	.21			.12		
<b>Step 3</b>		.15	.02		.08	.02
Openness	.18			.13		
<b>Step 4</b>		.17	.02		.12	.04*
LGD Problem solving	.14			.21*		

N = 93. \* $p < .05$ . \*\* $p < .01$ . Cohort was coded into four values (1-2-3-4) in accordance with the year of assessment (1998, 1999, 2000, 2001).

value of the LGD rating on problem solving, we entered the controls in the first step, cognitive ability test scores in the second step, and the Openness rating in the third step because Openness is conceptually related to problem solving (see Lievens et al., 2006). As a common thread running through each of these hierarchical regressions, the LDG rating on a specific dimension explained incremental variance over the conceptually related personality rating (and cognitive ability test score), implying that the added value of the LGD rating is due to the LGD method (and not because another construct was assessed). In particular, Hypotheses 2a and 2b were fully supported. There was partial support for Hypothesis 2c: The additional explained variance was observed for number of promotions ( $\Delta R^2 = .04$ ,  $p < .05$ ) but not for promotion speed ( $\Delta R^2 = .02$ , ns).

Some limitations should be acknowledged. A first limitation refers to the relatively small sample of officers and to the fact that the predictor methods were administered for research purposes. Clearly, these results should be replicated in larger samples and in actual selection con-

texts. Second, promotion criteria were used. Although these promotion criteria have often been used in assessment center research (e.g., Gaugler et al., 1987; Jansen & Stoop, 2001) and we tried to control for two key potentially contaminating variables, they might be influenced by other aspects such as political skills (Blickle et al., 2008). Third, data concerning candidates who left the institution ( $n = 32$ ) were not available. It would be interesting to investigate the reasons for their decision or their current job performance but this information was difficult to obtain for reasons of confidentiality. We did run an additional logistic regression analysis in which we used dropout (no = 0 or yes = 1) as the dependent variable and all the other predictors and controls as independent variables. Results showed that personality factors (and particularly Openness score) add a significant contribution in predicting candidate dropout ( $\Delta Chi^2 = 25.13$ ,  $p < .01$ ). Moreover, LGD dimension scores accounted for a significant additional portion of the variance in predicting candidate dropout ( $\Delta Chi^2 = 69.19$ ,  $p < .01$ ). More information on this additional analysis is available from the first author.

Despite these caveats, the key result of this study is that it confirms the utility of organizing LGDs. We found that ratings in LGDs provide a unique contribution to the prediction of promotion (measured on average 6.5 years later) in high-level executive positions beyond cognitive ability and personality.

## Acknowledgement

Funding for this study came from the French Navy (Marine Nationale).

## Note

1. The positive albeit insignificant correlation between the 'problem solving' dimension and GMA was not expected because this dimension is usually considered to be strongly related to GMA (e.g., Furnham et al., 2008).

## References

- Ansbacher, H. L. (1951). The history of the leadership group discussion technique. *Psychological Bulletin*, 48, 383–391. doi: 10.1037/h0055021
- Arthur, W., Day, E. A., Mcnelly, T. L., & Edens, P. S. (2003). A meta-analysis of the criterion-related validity of assessment center dimensions. *Personnel Psychology*, 56, 125–153. doi: 10.1111/j.1744-6570.2003.tb00146.x
- Arthur, W., & Villado, A. J. (2008). The importance of distinguishing between constructs and methods when comparing predictors in personnel selection research and practice. *The Journal of Applied Psychology*, 93, 435–442. doi:10.1037/0021-9010.93.2.435
- Barry, B., & Stewart, G. L. (1997). Composition, process, and performance in self-managed groups: The role of personality.

- The Journal of Applied Psychology*, 82, 62–78. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/9119798>
- Bass, B. M. (1954). The leadership group discussion. *Psychological Bulletin*, 51, 271–282. doi:10.1037/h0056881
- Blickle, G., Meurs, J. a., Zettler, I., Solga, J., Noethen, D., Kramer, J., & Ferris, G. R. (2008). Personality, political skill, and job performance. *Journal of Vocational Behavior*, 72, 377–387. doi:10.1016/j.jvb.2007.11.008
- Bono, J. E., & Judge, T. a. (2004). Personality and transformational and transactional leadership: A meta-analysis. *The Journal of Applied Psychology*, 89, 901–910. doi:10.1037/0021-9010.89.5.901
- Borteyrou, X., Bruchon-Schweitzer, M., Rasclé, N., & Collomb, P. (2006). Construction et validation d'une épreuve de groupe élaborée dans le cadre d'un centre d'évaluation pour des officiers de marine. *L'Orientation Scolaire et Professionnelle*, 35, 335–353. doi:10.4000/osp.1190
- Callinan, M., & Robertson, I. T. (2000). Work sample testing. *International Journal of Selection and Assessment*, 8, 248–260. doi:10.1111/1468-2389.00154
- Cattell, R. B. (1974). *Manuel: Test de facteur "g" de R. B. Cattell*. Paris: ECPA.
- Colom, R., & García-López, O. (2003). Secular gains in fluid intelligence: Evidence from the Culture-Fair intelligence test. *Journal of Biosocial Science*, 35, 33–39. doi:10.1017/S0021932003000336
- Costa, P. T., McCrae, R. R., & Rolland, J. P. (1998). *NEO PI-R inventaire de personnalité révisé*. Paris: Les Editions du Centre de Psychologie Appliquée.
- Eurich, T. L., Krause, D. E., Cigularov, K., & Thornton, G. C. (2009). Assessment centers: Current practices in the United States. *Journal of Business and Psychology*, 24, 387–407. doi:10.1007/s10869-009-9123-3
- Flanagan, J. C. (1954). The critical incident technique. *Psychological Bulletin*, 51, 327–358. doi:10.1037/h0061470
- Furnham, A., Taylor, J., & Chamorro-premuzic, T. (2008). Personality and intelligence correlates of assessment center exercises. *Individual Differences Research*, 6, 181–192.
- Gaugler, B. B., Rosenthal, D. B., Thornton, G. C., & Bentson, C. (1987). Meta-analysis of assessment center validity. *Journal of Applied Psychology*, 72, 493–511. doi:10.1037/0021-9010.72.3.493
- Gleason, W. J. (1957). Predicting army leadership ability by modified leaderless group discussion. *Journal of Applied Psychology*, 41, 231–235. doi:10.1037/h0044983
- Goffin, R. D., Rothstein, M. G., & Johnston, N. G. (1996). Personality testing and the assessment center: Incremental validity for managerial selection. *Journal of Applied Psychology*, 81, 746–756. doi:10.1037/0021-9010.81.6.746
- Griffin, B., & Hesketh, B. (2004). Why openness to experience is not a good predictor of job performance. *International Journal of Selection and Assessment*, 12, 243–251. doi:10.1111/j.0965-075X.2004.278\_1.x
- Hoffman, B. J., Melchers, K. G., Blair, C. a., Kleinmann, M., & Ladd, R. T. (2011). Exercises and dimensions are the currency of assessment centers. *Personnel Psychology*, 64, 351–395. doi:10.1111/j.1744-6570.2011.01213.x
- Hogan, R., & Shelton, D. (1998). A socioanalytic perspective on job performance. *Human Performance*, 11, 129–144. doi:10.1207/s15327043hup1102&3\_2
- Howard, A., & Bray, D. W. (1988). *Managerial lives in transition: Advancing and changing times*. New York: The Guilford Press.
- Hurtz, G. M., & Donovan, J. J. (2000). Personality and job performance: the Big Five revisited. *The Journal of Applied Psychology*, 85, 869–879. doi:10.1037/0021-9010.85.6.869
- Jansen, P. G., & Stoop, B. A. (2001). The dynamics of assessment center validity: Results of a 7-year study. *The Journal of Applied Psychology*, 86, 741–753. doi:10.1037/0021-9010.86.4.741
- Jansen, P. G., & Vinkenburg, C. J. (2006). Predicting management career success from assessment center data: A longitudinal study. *Journal of Vocational Behavior*, 68, 253–266. doi:10.1016/j.jvb.2005.07.004
- Jones, A. (1981). Inter-rater reliability in the assessment of group exercises at a UK assessment centre. *Journal of Occupational Psychology*, 54, 79–86. doi:10.1111/j.2044-8325.1981.tb00046.x
- Judge, T. a., Higgins, C. a., Thoresen, C. J., & Barrick, M. R. (1999). The Big Five personality traits, general mental ability, and career success across the life span. *Personnel Psychology*, 52, 621–652. doi:10.1111/j.1744-6570.1999.tb00174.x
- Kaess, W. a., Witryol, S. L., & Nolan, R. E. (1961). Reliability, sex differences, and validity in the leaderless group discussion technique. *Journal of Applied Psychology*, 45, 345–350. doi:10.1037/h0043681
- Kasten, R., Dayan, K., & Fox, S. (2002). Entry-level police candidate assessment center: An efficient tool or a hammer to kill a fly? *Personnel Psychology*, 55, 827–849. doi:10.1111/j.1744-6570.2002.tb00131.x
- Kirchmeyer, C. (2002). Gender differences in managerial careers: Yesterday, today, and tomorrow. *Journal of Business Ethics*, 37, 5–24. doi:10.1023/A:1014721900246
- Kumazawa, R. (2010). Promotion speed and its effect on attrition of Navy-enlisted personnel: Addressing heterogeneity in high school credentials. *Applied Economics*, 42, 2563–2576. doi:10.1080/00036840801964450
- Kurecka, P. M., Austin, J. M., Johnson, W., & Mendoza, J. L. (1982). Full and errant coaching effects on assigned role leaderless group discussion performance. *Personnel Psychology*, 35, 805–812. doi:10.1111/j.1744-6570.1982.tb02223.x
- Lazear, E. P., & Rosen, S. (1981). Rank-order tournaments as optimum labor contracts. *Journal of Political Economy*, 89, 841. doi:10.1086/261010
- Lievens, F. (2001). Assessor training strategies and their effects on accuracy, interrater reliability, and discriminant validity. *The Journal of Applied Psychology*, 86, 255–264. doi:10.1037/0021-9010.86.2.255
- Lievens, F. (2002). An examination of the accuracy of slogans related to assessment centres. *Personnel Review*, 31, 86–102. doi:10.1108/00483480210412436
- Lievens, F., Chasteen, C. S., Day, E. A., & Christiansen, N. D. (2006). Large-scale investigation of the role of trait activation theory for understanding assessment center convergent and discriminant validity. *The Journal of Applied Psychology*, 91, 247–258. doi:10.1037/0021-9010.91.2.247
- Lievens, F., Harris, M. M., Van Keer, E., & Bisqueret, C. (2003). Predicting cross-cultural training performance: The validity of personality, cognitive ability, and dimensions measured by an assessment center and a behavior description interview. *Journal of Applied Psychology*, 88, 476–489. doi:10.1037/0021-9010.88.3.476
- McHenry, J. J., Hough, L. M., Toquam, J. L., Hanson, M. A., & Ashworth, S. (1990). Project a validity results: The relationship

- between predictor and criterion domains. *Personnel Psychology*, 43, 335–354. doi:10.1111/j.1744-6570.1990.tb01562.x
- Meriac, J. P., Hoffman, B. J., Woehr, D. J., & Fleisher, M. S. (2008). Further evidence for the validity of assessment center dimensions: A meta-analysis of the incremental criterion-related validity of dimension ratings. *Journal of Applied Psychology*, 93, 1042–1052. doi:10.1037/0021-9010.93.5.1042
- Mussel, P., Winter, C., Gelléri, P., & Schuler, H. (2011). Explicating the openness to experience construct and its subdimensions and facets in a work setting. *International Journal of Selection and Assessment*, 19, 145–156. doi:10.1111/j.1468-2389.2011.00542.x
- Ng, T., Eby, L., Sorensen, K., & Feldman, D. (2005). Predictors of objective and subjective career success: A meta-analysis. *Personnel Psychology*, 58, 367–408. doi:10.1111/j.1744-6570.2005.00515.x
- Petty, M. (1974). A multivariate analysis of the effects of experience and training upon performance in a leaderless group discussion. *Personnel Psychology*, 27, 271–282. doi:10.1111/j.1744-6570.1974.tb01533.x
- Ree, M. J., & Earles, J. a. (1991). Predicting training success: Not much more than g. *Personnel Psychology*, 44, 321–332. doi:10.1111/j.1744-6570.1991.tb00961.x
- Robertson, I. T., & Kandola, R. S. (1982). Work sample tests: Validity, adverse impact and applicant reaction. *Journal of Occupational and Organizational Psychology*, 55, 171–183.
- Salgado, J. (1998). Big Five personality dimensions and job performance in army and civil occupations: A European perspective. *Human Performance*, 11, 271–288. doi:10.1207/s15327043hup1102&3\_8
- Salgado, J. (2003). Predicting job performance using FFM and non-FFM personality measures. *Journal of Occupational and Organizational Psychology*, 76, 323–346. doi:10.1348/096317903769647201
- 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, 735–746. doi:10.1037//0021-9010.87.4.735
- Schmidt, F. L. (2002). The role of general cognitive ability and job performance: Why there cannot be a debate. *Human Performance*, 15, 187–210. doi:10.1207/S15327043HUP1501&02\_12
- Schmidt, F. L., & Hunter, J. E. (1998). The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. *Psychological Bulletin*, 124, 262–274. doi:10.1037/0033-2909.124.2.262
- Seibert, S. E., & Kraimer, M. L. (2001). The five-factor model of personality and career success. *Journal of Vocational Behavior*, 58, 1–21. doi:DOI 10.1006/jvbe.2000.1757
- Shipilov, A., Labianca, G., Kalnysh, V., & Kalnysh, Y. (2014). Network-building behavioral tendencies, range, and promotion speed. *Social Networks*, 39, 71–83. doi:10.1016/j.socnet.2014.03.006
- Thornton, G. C., III, & Mueller-Hanson, R. A. (2004). *Developing organizational simulations: A guide for practitioners and students*. Mahwah, NJ: Erlbaum.
- Tippins, N. T., & Adler, S. (2011). *Technology-enhanced assessment of talent*. San Francisco, CA: Jossey-Bass.
- Waldman, D. a, Atwater, L. E., & Davidson, R. a. (2004). The role of individualism and the Five-Factor Model in the prediction of performance in a leaderless group discussion. *Journal of Personality*, 72, 1–28. doi:10.1111/j.0022-3506.2004.00254.x
- Weekley, J. A., Hawkes, B., Guenole, N., & Ployhart, R. E. (2015). Low-fidelity simulations. *Psychology and Organizational Behavior*, 2, 295–322. doi:10.1146/annurev-orgpsych-032414-111304