A Test of the Generalizability of a Recently Suggested Conceptual Model for Assessment Center Ratings

Natalia Merkulova
Universität Zürich, Switzerland

Klaus G. Melchers
Universität Ulm, Germany

Martin Kleinmann
Universität Zürich, Switzerland

Hubert Annen and Tibor Szvircsev Tresch
Militärrakademie an der ETH Zürich, Switzerland

Author Note
Natalia Merkulova and Martin Kleinmann, Psychologisches Institut, Universität Zürich; Klaus Melchers, Institut für Psychologie und Pädagogik, Universität Ulm; Hubert Annen and Tibor Szvircsev Tresch, Militärrakademie an der ETH Zürich, Switzerland.

Correspondence concerning this article should be addressed to Natalia Merkulova, Psychologisches Institut, Universität Zürich, Binzmühlestrasse 14/12, CH-8050 Zürich, Switzerland, e-mail: nmerkulova@sunrise.ch; or to Klaus Melchers, Institut für Psychologie und Pädagogik, Universität Ulm, Albert-Einstein-Allee 47, D-89069 Ulm, Germany, e-mail: klaus.melchers@uni-ulm.de.

Keywords: assessment center, construct-related validity, criterion-related validity, personnel selection
Abstract

The present study replicated and extended research concerning a recently suggested conceptual model of the underlying factors of dimension ratings in assessment centers (ACs) proposed by Hoffman, Melchers, Blair, Kleinmann, and Ladd (2011) that includes broad dimension factors, exercise factors, and a general performance factor. We evaluated the criterion-related validity of these different components and expanded their nomological network. Results showed that all components (i.e., broad dimensions, exercises, general performance) were significant predictors of training performance. Furthermore, broad dimensions showed incremental validity beyond exercises and general performance. Finally, relationships between the AC factors and individual differences constructs (e.g., Big Five, core self-evaluations, positive and negative affectivity) supported the construct-related validity of broad dimensions and provided further insights in the nature of the different AC components.

Keywords: assessment center, construct-related validity, criterion-related validity, personnel selection
A test of the generalizability of a recently suggested conceptual model for assessment center ratings

Although assessment center (AC) users have traditionally relied on dimension scores for the design and interpretation of ACs, there still remains a lack of agreement whether dimensions are really relevant building blocks of ACs (Lance, 2008). Specifically, evidence suggests that exercises explain more variance in post-exercise dimension ratings (PEDRs) from ACs than dimensions (see Melchers, Henggeler, & Kleinmann, 2007, or Woehr & Arthur, 2003, for meta-analytic evidence). Moreover, confirmatory factor analyses (CFAs) conducted on the basis of PEDRs from ACs found that dimensions rarely emerge as latent factors whereas exercises do (Lance, Lambert, Gewin, Lievens, & Conway, 2004). Additionally, instead of separate dimensions, a general performance factor emerges regularly that seems to underlie performance across all PEDRs in an AC (Lance, Lambert et al., 2004). Taken together, these results were long considered problematic for the construct-related validity of AC dimensions and even led some researchers to suggest that dimensions are not the building blocks of ACs (Jackson, Barney, Stillman, & Kirkley, 2007; Jackson, 2012; Lance, 2008).

However, recently Hoffman, Melchers, Blair, Kleinmann, and Ladd (2011) proposed a new approach to modeling AC performance and found evidence that seems to be much more promising concerning the construct-related validity of AC dimension ratings. Specifically, they suggested that a mixed-model architecture with exercise factors, a general performance factor, and broad dimension factors that group together sets of similar dimensions represents a more appropriate structure for AC ratings. Furthermore, they presented evidence from four samples that such an underlying structure provided a close fit to AC data and also a significantly better fit than previously suggested models. In light of past findings, the support for AC dimensions using this new structure was promising. Yet, given the persistent and
consistent results questioning the construct-related validity of dimension ratings (Lance, 2008; Lance, Lambert et al., 2004), it is important to determine whether this new structure of AC ratings is replicable and generalizes to other assessment centers.

In addition, more knowledge concerning the correlates of the different latent AC factors is needed for at least three reasons. First, in Hoffman et al.’s (2011) study, criterion data were available for only two samples. We therefore wanted to evaluate the relationships between the different latent factors and a criterion that was not considered by Hoffman et al., namely training performance. Second, in contrast to previous studies that only modeled a general performance factor but no dimensions factors (Lance, Lambert, et al., 2004), Hoffman et al. found no support for the criterion-related validity of the general performance factor. Thus, their results question the relevance of this factor for predicting performance criteria when broad dimension factors are modeled. Therefore, it is important to further explore the relevance of the general performance factor for criterion-related validity and hereby also to address recent calls for research regarding this factor (Kuncel & Sackett, 2014; Thornton & Gibbons, 2009). Finally, knowledge concerning the nomological network of the new latent AC factors should be expanded. As Hoffman et al. were only able to consider few individual difference variables for only two of the samples, more evidence with regard to the relationships of the latent AC factors with additional work-related dispositions is needed.

To address these gaps in our knowledge, the first aim of the present research was to evaluate the generalizability of the recently suggested structure of AC ratings – especially concerning the existence of broad dimension factors – in a different setting. Our second aim was to further explore the criterion-related validity of the different latent AC factors and to test whether all three sources of variance contribute to the criterion-related validity of ACs. Finally, the third aim was to expand the nomological network of the different AC factors by involving a broad range of work-related individual difference variables.
Previous Research

Internal Structure of Assessment Center Ratings

Before we describe the conceptual structure recently suggested by Hofman et al. (2011), we will briefly outline the models that were regularly tested in prior AC research using CFAs. Specifically, based on different conceptualizations of ACs, previous research has repeatedly tested the following four models of the underlying structure of PEDRs from ACs.

Model 1 (J-dimensions, 0-exercises) postulates that there are no situational influences to performance in ACs and that only dimensions account for variance in AC ratings. This means that a latent factor represents each measured dimension, but that the specific exercises should not be of importance. This model usually fails to converge to an admissible solution (Lievens & Conway, 2001; Lance, Lambert et al., 2004). Model 2 (J-dimensions, K-exercises) presupposes that correlated exercises as well as correlated dimensions represent the internal structure of an AC. This means that the variance in PEDRs should be attributable on one hand to the different dimensions that were rated and on the other hand to the different exercises. Empirically, however, this model also shows systematic problems of admissibility and convergence in CFAs (Lance, Lambert et al., 2004; Lance et al., 2000). Model 3 (0-dimensions, K-exercises) only includes multiple correlated exercise factors but no dimension factors, assuming that AC performance is solely triggered by situational cues and that it is inconsistent across exercises. This model commonly converges to an admissible solution. Finally, Model 4 (1-general performance factor, K-exercises) assumes that some aspect of performance in an AC is consistent across exercises and may be represented by a single general performance factor in addition to the different exercise factors. A large scale quantitative review by Lance, Lambert et al. (2004) suggests that this model is usually admissible and often provides a better fit in comparison to the previously mentioned models.

Taken together, previous CFA-based AC research has rarely found support for models that involve the targeted AC dimensions. This condition has led some researchers to suggest
that ACs mainly measure cross-situationally specific performance (Jackson, Stillman, & Atkins, 2005; Lance, 2008; Lance et al., 2000). This view was supported by findings that showed that exercise factors not only reflect variance in PEDRs but are also related to job performance (Lance, Foster, Nemeth, Gentry, & Drollinger, 2007; Lance et al., 2000).

However, findings from different domains of performance assessment advocate that both different targeted dimensions and different situations explain variance in evaluations of ratees’ performance (Hoffman, 2012; Melchers, Wirz, & Kleinmann, 2012). For example, in structured interviews dimensions and different kinds of interview questions contribute to variance in ratings of interviewees’ performance (Klehe, König, Richter, Kleinmann, & Melchers, 2008; Van Iddekinge, Raymark, Eidson, & Attenweiler, 2004). Similarly, variance in multisource feedback ratings can be traced to dimensions and different rater sources (i.e., subordinates vs. peers vs. supervisors; e.g., Hoffman, Lance, Bynum, & Gentry, 2010).

Moreover, theory and measurement of leadership and managerial performance has regularly used broad categories of narrow manifest dimensions to describe performance (Hoffman, 2012). For example, conceptual taxonomies of managerial performance (e.g., Borman & Brush, 1993), theoretical and empirical models of organizational citizenship behavior (Hoffman, Blair, Meriac, & Woehr, 2007; Smith, Organ, & Near, 1983), measures of managerial behavior and skills (McCauley, Lombardo, & Usher, 1989), and also investigations of the internal structure of multisource performance ratings (Hoffman et al., 2010) usually model the structure of performance as narrow manifest dimensions that load on broader dimension factors.

In addition, there are reasons to believe that AC dimensions fail to emerge in CFAs because of empirical under-identification problems (Brannick & Spector, 1990; Kenny & Kashy, 1992). In particular, if the manifest dimensions are too similar – which is often the case with some AC dimensions (Arthur, Day, McNelly, & Edens, 2003) – then this might also lead to problems of inadmissibility in CFAs (Hoffman, 2012). Furthermore, findings from
CFA research suggest that models with a low ratio of indicators vs. latent factors regularly result in improper solutions while increasing indicator-factor ratio often helps to obtain admissible solutions for CFA models (Tomas, Hontangas, & Oliver, 2000). This means that if AC dimensions are modeled with a greater number of indicators there is a greater chance of achieving admissible solutions (Monahan, Hoffman, Lance, Jackson, & Foster, 2013).

For these reasons, Hoffman et al. (2011) proposed to group together similar AC dimensions that represent broader constructs and thus introduced a new model of AC ratings. Accordingly, their model (Model 5) includes broad dimension factors in addition to exercise factors and a general performance factor (J-broad dimensions, K-exercises, and L-general performance factor). The broad dimensions in this model are conceptualized as latent factors on which ratings of similar but separate narrow dimensions load together as indicators of these latent factors. Hoffman et al. tested this model and found that it provided a significantly better fit of the structure of AC ratings than all formerly suggested models. Furthermore, since none of the traditional dimension models converged to a proper solution in their study, this new model also appeared to be a superior model when it came to modeling AC dimensions. Thus, when dimensions were modeled in a way that took the similarity between them into account, then it was possible to find evidence for dimension factors in ACs.

However, despite these initial promising results, the generalizability of this new structure of AC ratings still needs to be explored more extensively in other contexts. Given that each AC administration is unique and different in terms of design characteristics, like the kind and number of dimensions rated, type of participants and assessors, etc. (Woehr & Arthur, 2003), there is a need for additional evidence that supports the structure suggested by Hoffman et al. (2011). Moreover, given the large body of research that failed to find supportive evidence for AC dimensions in PEDRs (Lance et al., 2000; Lance et al., 2007; Melchers et al., 2007; Woehr & Arthur, 2003), more research regarding the proposed new model and the nature of the new latent AC factors is needed. Therefore, the first aim of the
present study was to test the generalizability of the proposed new structure of AC ratings in a different AC setting.

**Criterion-Related Validity of the Latent AC Factors**

For two samples for which criterion data were available, Hoffman et al. (2011) found that broad dimension and exercise factors contributed to the criterion-related validity of the AC, while the general performance factor did not. In one sample, broad dimensions correlated significantly with salary growth, while neither the general performance factor nor exercise factors did. In the other sample, broad dimensions and exercise factors were significantly related to supervisor ratings of job performance, while again the general performance factor was not. Finally, in both samples, broad dimensions explained incremental variance in job performance over and above exercise factors and the general performance factor.

These findings somewhat challenge results from previous research that found that the general performance factor in Model 4 (1-general factor, K-exercises) significantly contributed to criterion variance (Lance et al., 2000; Lance, Foster, Gentry, & Thoresen, 2004). A possible reason for the diverging findings concerning the criterion-related validity of the general performance factor is that previous results by Lance and colleagues reflect the effects of unmodeled broad dimension factors and not the general performance factor per se. Said differently, it is possible that the variance in prior studies that was due to broad dimensions might erroneously have been attributed to the general performance factor when no broad dimension factors were modeled.

However, before the relevance of the general performance factor for AC criterion-related validity can be disputed in general, its contribution needs to be evaluated in further studies. First, it seems surprising that the general performance factor – a factor that reflects consistent aspects of AC performance across all exercises and dimensions – is not a significant predictor of actual performance criteria. Second, given support for a general performance factor in job performance ratings (Viswesvaran, Schmidt, & Ones, 2005) and
given that ACs are designed to measure valid samples of performance, some correspondence between the general performance factor in an AC and performance on the job is expected. Third, given that there is disagreement on whether the general performance factor reflects potential rater bias (see, for example, the discussion of this in Woehr, Meriac, & Bowler, 2012) or something very different like a general ability that is effective across a variety of situations (Thornton & Gibbons, 2009), there is a need to gain further insight concerning the nature of this factor.

Thus, to explore whether Hoffman et al.’s findings generalize to other contexts, there is a need to evaluate the criterion-related validity of the new latent factors and also to include additional performance criteria. One such criterion is training performance. Training performance has previously been used to determine the criterion-related validity of ACs (Schmidt & Hunter, 1998), however, in that research, the validity of the overall assessment rating (OAR), but none of the different latent factors was considered. Therefore, we used training performance in our study.

Accordingly, the second aim of the present research was to explore the criterion-related validity of the different latent AC factors more extensively. In doing so, we wanted to examine the relevance of the general performance factor for criterion-related validity. In addition, we intended to expand previous findings by including additional performance criteria, like training performance.

**Nomological Network of AC Factors**

Another approach to determining the construct-related validity of the proposed underlying constructs of a selection procedure is to examine their nomological network, that means, their relationships with relevant external variables (American Educational Research Association, 2004). Past research that studied the nomological network of AC ratings involved the OAR (e.g., Collins et al., 2003), final dimension ratings (Dilchert & Ones, 2009; Meriac, Hoffman, & Woehr, 2014; Shore, Thornton, & Shore, 1990), overall exercise ratings
(Hoffman, Monahan, Lance, & Sutton, in press), or latent exercise factors together with the latent general performance factor (Lance et al., 2007) and their relationships with the Big Five and general mental ability (GMA). However, this research allows only limited inferences with regard to the construct-related validity of the different latent components that constitute AC performance, namely dimensions, exercises and general performance (Hoffman et al., 2011) relative to each other.

Moreover, although the Big Five and GMA have been useful as a framework to propose relationships between individual differences and AC factors, further relevant information can be gained from studying additional work-related constructs, such as, for example, the different core self-evaluations constructs and trait affectivity. The core self-evaluations constructs are important work-related variables because of their relationships with job satisfaction, motivation, and job performance (see Judge & Bono, 2001, for meta-analytic evidence). Moreover, positive and negative affectivity are also variables that are important for job performance because of their relationships with task performance and organizational citizenship behaviors (see Kaplan, Bradley, Luchman, & Haynes, 2009, for meta-analytic evidence). However, to our knowledge information on how the different components of AC performance relate to these work-related individual difference variables is still missing.

To address the limited knowledge concerning the nomological network of the different latent factors of AC performance, the final aim of the present study was to further explore the correlates of the different latent AC factors. On the one hand, we considered the relationships of these factors with individual difference constructs that were already included in previous studies, such as the Big Five factors and GMA. On the other hand, we also considered additional constructs that were not captured by previous AC research, namely the different core self-evaluations constructs as well as positive and negative affectivity.

We will now discuss previous findings and our assumptions concerning the relationships of the broad dimension factors, the general performance factor, and the exercise
factors with individual differences. Hoffman et al.’s (2011) initial study provided some preliminary results concerning individual differences and broad dimension factors. In their study, general mental ability (GMA) was partially related to a broad conceptual/administrative skills factor, conscientiousness was related to an interpersonal skills factor, and dominance was correlated with a leadership factor. These findings can be seen as initial support for the construct-related validity of broad dimension factors.

In the present study, we expect that the broad dimension factors will be related to a broad range of individual difference variables. First, as suggested by Thornton and Rupp (2006), some dimensions may be more “cognitively loaded” than others and, thus, should be related to GMA. For example, findings from a study with final dimension ratings suggest that dimensions that are associated with individuals’ administrative skills are more strongly related to GMA (Shore et al., 1990). This leads to the following hypothesis:

Hypothesis 1: Broad dimensions that are cognitively oriented will be positively related to GMA.

Second, based on studies with final dimension ratings (e.g., Dilchert & Ones, 2009; Lievens, Chasteen, Day, & Christiansen, 2006; Meriac et al., 2014; Shore et al., 1990) and findings concerning broad dimensions by Hoffman et al. (2011), we expect that the broad dimension factors will be related to the Big Five. Specifically, previous results suggest that conscientious individuals perform better during social interaction tasks due to their responsibility and reliability (Hoffman et al. 2011; LePine & Van Dyne, 2001; Meriac et al., 2014). In addition, extraverted individuals who are more sociable, confident, and enthusiastic should be more effective when interacting with others (LePine & Van Dyne, 2001; Meriac et al., 2014). Moreover, agreeable individuals who are more cooperative and friendly should also approach others more easily in interactive tasks (LePine & Van Dyne, 2001). Therefore, we expect that that conscientiousness, extraversion, and agreeableness will be related to dimensions associated with interpersonal skills. Furthermore, since extraversion is also
associated with assertiveness (Barrick & Mount, 1991) and leadership behaviors (Judge, Bono, Ilies, & Gerhardt, 2002), we expect that extraversion will be related to broad dimensions associated with drive and with leadership:

\textit{Hypothesis 2.1:} Broad dimensions that are associated with interpersonal skills will be positively related to agreeableness, conscientiousness, and extraversion.

\textit{Hypothesis 2.2:} Broad dimensions that are associated with drive and with leadership will be positively related to extraversion.

Third, we expect that the broad dimension factors will be related to the different core self-evaluations constructs. Based on assumptions by Judge, Erez, and Bono (1998), we believe that individuals who are high on core self-evaluations have a positive self-concept. As a positive self-concept is important in approaching others successfully (Judge et al., 1998), these individuals should receive better ratings in interactional tasks. Thus, we expect that the core self-evaluations constructs will be related to interpersonally-oriented dimensions. Furthermore, it has also been proposed that the core self-evaluation traits are indicative of beliefs about one’s self-regulatory and behavioral capacities (Johnson, Rosen, & Levy, 2008). These beliefs include the controllability of one’s actions in the environment, anticipated success as a result of a change in one’s actions and environment, as well as setting positive and motivating goals. We therefore believe that individuals who are high in core self-evaluations should be more motivated to take initiative, to seek responsibility, or to compete with others and, thus, these individuals should perform better with regard to dimensions that involve such behaviors. Therefore, we assume that the core self-evaluations constructs will also be related to broad dimensions that are associated with behaviors that are indicative of drive as well as behaviors that are indicative of leadership.
Hypothesis 3: Broad dimensions that are associated with interpersonal skills, drive, and leadership skills will be positively related to the different core self-evaluations constructs.

Fourth, we suggest that broad dimensions that are indicative of interpersonal behaviors, like interpersonal skills and leadership skills, will be related to trait affectivity. Individuals who have a tendency to experience more negative feelings such as fear, distress, and hostility, and who also tend to experience their environment in a negative way (Watson, Wiese, Vaidya, & Tellegen, 1999), should be less inclined to positively interact with others and may even be avoided by others (Judge et al., 1998). Therefore, trait negative affectivity should be negatively related to dimensions that are associated with interpersonal behavior, like interpersonal skills or leadership skills. Furthermore, individuals who are predisposed to experience positive moods (Watson et al., 1999) should be more proactive and engaged during interpersonal contacts and thus they should be more successful in shaping their interactions with others. Therefore, we suggest that trait positive affectivity will be positively related to dimensions associated with interpersonal skills and also with dimensions associated with leadership skills.

Hypothesis 4: Broad dimensions that are associated with interpersonal skills and leadership skills will be positively related to positive affectivity and negatively to negative affectivity.

Similar to the criterion-related validity of the general performance factor, findings concerning the nomological embedment of the general performance factor seem to be inconsistent. In Hoffman et al.’s (2011) studies, general performance was related to GMA in only one sample and was correlated with conscientiousness in both samples. In contrast, in earlier studies (Lance et al., 2000; Lance et al., 2007), general performance was not related to GMA, but was found to be significantly and moderately related to conscientiousness, emotional stability, and openness (Lance et al., 2007). However, in this case, it is also
possible that the absence of broad dimension factors in the models tested in the earlier studies led to a misinterpretation of correlates of the general performance factor. Said differently, variance due to individual differences that is probably characteristic for broad dimension factors might erroneously have been attributed to the general performance factor.

In the present study, based on previous findings (Collins et al., 2003; Hoffman et al., 2011; Lance et al., 2007), we suggest that the general performance factor is related to GMA. Since GMA is a “very general” capability that is considered critical in various areas of life (Gottfredson, 1997, p. 13) and since meta-analytic results also suggest that GMA is related to AC performance (Collins et al., 2003), we believe that GMA should be related to AC performance throughout the entire AC or, in other words, to the general performance factor.

_Hypothesis 5:_ The general performance factor will be positively related to GMA.

Finally, information concerning the relationships between exercise factors and individual difference constructs is still rather limited. Some findings suggest that exercises measure cross-situationally specific behavior, so that exercise factors should not be related to individual differences (Lance et al., 2007). In line with this, only very few correlates of exercise factors were found in Hoffman et al.’s (2011) study. Specifically, in one of their samples, Hoffman et al. found that extraversion and conscientiousness were marginally related to an exercise factor associated with a role-play concerning a one-on-one meeting. Furthermore, GMA was related to a leaderless group discussion factor in their study. However, these correlations were much lower than the relationships between the broad dimension factors and individual differences and no other individual difference correlates were found for the remaining exercise factors. Based on these findings (Hoffman et al., 2011; Lance et al. 2007), we assume that the exercise factors represent rather situationally-specific behavior. This means that we do not expect meaningful relationships between these factors and individual difference variables.
Method

Sample

The sample consisted of 936 participants of an assessment center used to select career officers for the Swiss Armed Forces, of which 18 were females. The applicants were untenured reserve officers who were interested in becoming full-time military career officers. Those who passed the AC were granted training at the Swiss military training academy before they were employed on the basis of tenured contracts. The mean age of the candidates was 27 years, with a range from 19 to 43. The candidates’ educational backgrounds varied from apprenticeship to university degree.

Procedure and Assessment Center

The AC was similar to other ACs that adhere to current standards (International Task Force on Assessment Center Guidelines, 2015) and was previously shown to have good validity for predicting future training and job performance, as well as career advancement (Gutknecht, Semmer, & Annen, 2005; Melchers & Annen, 2010).

The AC consisted of the following six exercises: a) a short oral presentation, in which candidates had to give a short talk about themselves and an appointed topic, b) a leaderless group discussion, in which candidates were assigned to enforce their own interests but to also represent the interests of the group, c) a motivational talk, that required participants to motivate a role player to perform an unpleasant task, to accept a situation, or to not give up in a problematic situation, d) a debate, in which participants were assigned to either a pro or a con group and then had to convince the other group of their position, e) a set of short cases that consisted of different problematic situations that could occur in everyday military life and required candidates to describe how they would react, f) a lecture on a topic of military pedagogy, which had to be prepared by the candidates during their free time between the AC exercises on the basis of provided materials.
Candidates was rated by two experienced assessors (a military officer trained in human resource management and a psychologist from the field of psychological assessment) who received at least one day of frame-of-reference training (Roch, Woehr, Mishra, & Kiesczczynska, 2012). Following each exercise in the AC, every candidate was rated on three to six dimensions with a four point scale ranging from 1 (= clearly failed to meet requirements) to 4 (= clearly exceeded requirements). These were personal attitude (e.g., self-confident manner, being able to deal with own emotions), achievement motivation (e.g., showing commitment and persistence), analysis and planning (e.g., structured and purposeful way of handling and explaining situations), dealing with conflicts (e.g., recognizing conflict potential, offering consensual solutions), influencing others (e.g., being able to convince/motivate others, to present good arguments), social contact (e.g., being able to work in a team, facing others with openness), and oral communication (e.g., being able to express oneself clearly). The targeted dimensions were somewhat different from those usually measured in ACs for the selection of managers. For example, dimensions like planning and organizing, decision making, and leadership were not measured in the AC. The reason for these slight differences is that planning and organizing and decision making were assessed before when the candidates completed another selection procedure to become militia officers. Furthermore, the targeted dimensions influencing others and analysis and planning comprised different aspects, usually associated with leadership skills.

Data were collected over a period of 17 years from 1993 until the end of 2009. The ACs took place three times a year and approximately 30 candidates took part in each administration. Over the years, questionnaires that assessed the individual differences constructs that were relevant for the present study were successively included in the AC. Specifically, GMA was measured starting from 1996, the Big Five were assessed starting from 2000, and positive and negative affectivity and core self-evaluations were included in 2003. The candidates received a questionnaire that assessed the individual differences
constructs in the introduction meeting of the AC and were asked to complete it as honestly as possible during their free time between the exercises. They were also told that the results of the personality questionnaires would solely be used for research purposes and that their responses in the personality tests would support the AC administrators and researchers in their attempt to ensure the high validity and quality of the AC.

**Criterion Variables**

On average, data on criteria were collected 3.17 years after the AC for candidates who completed a three-year Bachelor course and 1.30 years for candidates who completed a one-year training course (see below) at the military academy. Two different criteria were used, academic training performance and military training performance. Neither the faculty members at the military academy nor the candidates’ supervisors had knowledge of the candidates’ specific AC results. These results were provided only to the candidates and to the administrating officer of the human resources department of their arms unit. These administrating officers were not involved in the evaluation of the candidates’ academic or military performance in the different training courses.

*Academic training performance.* The first criterion was the candidates’ later academic performance at the military academy, and it was operationalized as the final course grade of the academic training course. This final course grade was available for a total of 500 candidates. Depending on their previous qualifications, candidates who succeeded in the AC either attended a three-year Bachelor course at the military academy (candidates who had A-levels; \( n = 312 \)), a one-year diploma course (candidates who already had a university degree; \( n = 109 \)) or a one-year Military School course (candidates who had completed an apprenticeship and were experienced and qualified militia officers; \( n = 79 \)). The final course grades comprised various grades from written exams during the military academic courses and the grade from the Bachelor’s thesis. Since participants’ from the Bachelor course were rated on a scale ranging from 1 to 6s and other participants were rated on a scale ranging from
1 to 5 in the other two courses (with higher grades always indicating better performance), the final grades were z-standardized within each course before the analyses. Before the z-standardization, the means and SDs for candidates from the different courses were 4.88 (SD = 0.37), and 3.38 (SD = 0.53), and 3.21 (SD = 0.37) for the three-year Bachelor course, the one-year diploma course, and the one-year Military School course respectively.

Military training performance. The candidates’ later military performance was used as the second criterion. This variable stemmed from the practical military training that the participants completed in the periods between their academic courses. Each year, the participants were usually evaluated by their direct military supervisors with regard to their general behavior as an officer (e.g., their attitude, motivation, and communication) as well as their performance in specific military training (tactics, staff procedures, and combat training). Furthermore, the supervisors also rated the participants’ overall performance on an additional single item. All performance ratings were made on a scale ranging from 1 to 5 with higher values reflecting better performance. For the present study, we used the mean of these ratings including the overall rating at the end of the military training (n = 514, M = 3.39, SD = 0.60).

Individual Difference Variables

GMA. As part of the AC, each participant completed three written cognitive ability tests. The tests measured verbal, numerical and abstract non-verbal reasoning. The first test measured candidates’ understanding of short but complex texts. In the second test, candidates had to analyze tables and graphs in order to answer questions related to each table or graph. In the third test, candidates had to derive the rule behind a series of four diagrams and then determine a corresponding fifth diagram that was missing. In the present study, the candidates’ mean across the three tests was used as an indicator of their GMA. The tests were purchased from an international consulting firm that had developed and pre-tested them. According to the manual, the internal consistencies were .75, .81, and .80 for the verbal test, numerical and reasoning tests, respectively. The tests were previously shown to have
(uncorrected) validities between .21 and .28 for predicting overall job performance (SHL, 2006).

**Big Five.** The Big Five personality traits were assessed with a shortened version of the minimal redundant scales (Ostendorf, 1990) by Schallberger and Venetz (1999). These scales measure personality traits with four paired adjectives each (e.g., neuroticism: emotionally stable – unstable, extraversion: sociable – withdrawn, conscientiousness: orderly - unordered, openness to experience: creative - uncreative, agreeableness: good natured – short-tempered). The paired adjectives were presented as two end-points of a scale from 1 to 6 and the candidates were required to indicate whether they were closer to one or the other of the paired adjectives. The internal consistencies of these scales were .60 for emotional stability, .74 for extraversion, .77 for conscientiousness, .75 for openness, and .43 for agreeableness (.54 after excluding 2 items).

**Positive and negative affectivity.** A shortened and modified version of the Positive and Negative Affect Scale (PANAS) originally developed by Watson, Clark, and Tellegen (1988) was used. This version was developed on the basis of the German translation of the PANAS by Krohne, Egloff, Kohlmann, and Tausch (1996) and was modified by Schallberger (2005) to measure these dimensions with eight bipolar items. In this eight-item version, four adjective pairs measured positive affectivity (e.g., bored – enthusiastic), while another four measured negative affectivity (e.g., relaxed – stressed). Similar to the Big Five scales used, candidates had to indicate on a scale from 1 to 6 which adjective corresponded with their personality to a greater extent. The internal consistencies of the scales were .79 for positive and .66 for negative affectivity.

**Locus of control and self-efficacy.** A questionnaire by Krampen (1991) was employed to measure locus of control (e.g., “If I have an accident or not, depends entirely on me and my behavior”) with eight items and self-efficacy (e.g., “Even in difficult situations I always come up with ideas about what can be done”) with four items. The items had to be rated on 6-point
scales with a range from 1 (\textit{completely false}) to 6 (\textit{completely true}) for locus of control and a similar range from 1 (\textit{completely false}) to 6 (\textit{completely true}) for self-efficacy. The internal consistency was .61 for locus of control and .69 for self-efficacy.

\textbf{Self-esteem.} Self-esteem was measured with Badura’s (1987) 10-item German translation of a scale by Rosenberg (1965). The items (e.g., “Sometimes I really feel worthless”) were rated on a 5-point scale ranging from 1 (\textit{completely false}) to 5 (\textit{completely true}). The internal consistency of this scale was .67.

\textbf{General Analytic Approach}

\textit{CFA analyses.} We used AMOS 18.0 (Arbuckle, 2009) to conduct confirmatory factor analyses to evaluate the five models described above. For models that included (either narrow or broad) dimensions, these dimensions were allowed to correlate with each other. Similarly, for models that included exercise factors, these factors were allowed to correlate with each other. However, dimension factors, exercise factors, and the general performance factor were conceptualized as being uncorrelated with each other. Also, uncorrelated error terms were assumed for the PEDRs.

Several criteria were used as indicators of the goodness of fit of a model. First, the model had to converge to a proper solution (i.e., minimization was successful, standardized model parameters did not exceed the absolute value of 1.00, etc.). Second, the models were evaluated according to the $\chi^2$ statistic, the standardized root mean squared residual (SRMSR), the root mean squared error of approximation (RMSEA), the comparative fit index (CFI), and the Tucker-Lewis index (TLI). Finally, $\Delta\chi^2$-tests were conducted to determine whether the improvement in model fit between competing nested models was significant.

\textit{Broad dimension factors.} To determine the underlying broad dimensions of the given AC, we tried different combinations of the measured dimensions by assuming that PEDRs of similar dimensions load on one latent factor. We classified the manifest dimensions into broad dimensions by taking their similarity with each other into account and also by considering
three previously proposed dimension taxonomies. The first taxonomy consisted of two
general broad categories by Shore et al. (1990) who suggested that PEDRs could be
differentiated into a performance style dimension, which should be more strongly related to
ability measures, and an interpersonal style dimension. The second taxonomy consisted of
four broad higher-order categories that were determined on the basis of sorting 187
dimensions of managerial performance into categories (Borman & Brush, 1993). These
categories were interpersonal dealings and communication, leadership and supervision,
technical activities and the “mechanics of management” (e.g., planning and organizing,
problem solving etc.), and useful personal behavior and skills (e.g., persistence, resilience,
organizational commitment etc.). In a third approach, Arthur et al. (2003) collected a large
number of dimensions from previous AC research and derived a set of seven categories to
categorize nearly all dimensions. These seven dimensions were consideration/awareness of
others, communication, drive, influencing others, organizing and planning, and problem
solving.

Recent meta-analytic work by Meriac et al. (2014) analyzed whether the dimension
structure proposed by Arthur et al. (2003) is indeed empirically supported in meta-analytically
derived correlational patterns of final dimension ratings. Their results suggest that the
structure of AC dimensions is best represented by a parsimonious model involving three
overarching dimension factors that comprise dimensions related to drive, to interpersonal
skills, and to administrative skills. These results were also supported by the internal structure
of broad dimensions found in the studies by Hoffman et al. (2011). Since not all manifest
dimensions in our study were easily classified into one of these mentioned broad categories,
we studied the behavioral anchors of the dimensions and correspondingly tried different
combinations of dimensions that made sense. We then chose the best solution in terms of
model admissibility and correspondence with existing taxonomies for our Model 5.
Relationships with external variables. After the best fitting model was identified, we estimated the relationships between the AC factors and external variables. Similarly to Hoffman et al. (2011) and Lance et al. (2000), we fixed the AC parameter values to the estimates from the CFA. Then we separately included the external variables to the model to estimate their correlations with the latent factors. For the personality variables, item level data were available so that separate items were used as indicators. The factor loadings of the items on the latent personality factor and their respective uniquenesses were then freely estimated in the model. For the variables and criteria for which only one indicator was available to us, namely GMA, academic training performance, and military training performance (supervisor ratings), we set the factor loadings to the square root of their reliability if known (GMA: $\alpha = .79$) or to the square root of their meta-analytic reliability estimates, namely .80 for academic training performance (Hülsheger, Maier, & Stumpp, 2007; Hunter & Hunter, 1984) and .86 for military training performance (supervisor ratings; Viswesvaran, Ones, & Schmidt, 1996) and the error variances to one minus their respective reliabilities.

Results

Confirmatory Factor Analyses

We first tested the traditional models of AC ratings with our data and then analyzed different conceptualizations of broad dimension factors. On the basis of previous research, we did not expect that models involving the manifest dimensions would converge to proper solutions. However, for the sake of completeness, we also included the traditional models in Table 1. Neither Model 1, which only consisted of the originally targeted narrow dimensions, nor Model 2, which contained the exercise factors in addition to the targeted dimensions, converged to a proper solution.

In line with prior research (Lance et al., 2000; Lance, Lambert et al., 2004), Model 4 (6-exercises, 1 general performance factor) converged to a proper solution and showed a good
fit according to all fit indices, $\chi^2(258) = 444.27$, $\chi^2/df = 1.72$, SRMSR = .026, RMSEA = .028, CFI = .978 and TLI = .972. Moreover, Model 4 demonstrated a significantly better fit than Model 3 (6-exercise factors only), which also converged to a proper solution, $\Delta \chi^2(26) = 134.20$, $p < .001$ (Table 1).

For Model 5, two conceptualizations with 6 exercise factors, 1 general performance factor, and either 2 (Model 5a) or 3 (Model 5b) broad dimensions yielded a good fit to the data and also showed good correspondence of the broad dimension factors with existing taxonomies (e.g., Arthur et al., 2003; Borman & Brush, 1993; Meriac et al., 2014; Shore et al., 1990). In the conceptualization with two broad dimension factors, the AC dimensions achievement motivation, analysis and planning, and influencing others loaded on a broad performance style factor, while personal attitude, oral communication, social contact, and dealing with conflicts loaded on a broad interpersonal style factor. The conceptualization with three broad dimension factors used the following specification of broad dimensions: interpersonal skills (personal attitude, dealing with conflicts, social contact, and oral communication), drive (achievement motivation), and strategic skills (influencing others, analysis and planning). As can be seen in Table 1, the different fit indices for both models reflected a good fit to the data: $\chi^2(232) = 345.71$, $\chi^2/df = 1.49$, SRMSR = .022, RMSEA = .023, CFI = .987 and TLI = .981 for Model 5a and $\chi^2(229) = 321.95$, $\chi^2/df = 1.41$, SRMSR = .023, RMSEA = .021, CFI = .989 and TLI = .984 for Model 5b.

In line with our expectations and as in the studies by Hoffman et al. (2011) and Meriac et al. (2014), the models with broad dimensions (Models 5a and 5b) had a significantly better fit than Model 4, both $\Delta \chi^2$s > 98.55, $\Delta df$s = 26 and 29, respectively, both $ps < .001$. In addition, the three-broad dimensions model (Model 5b) had a significantly better fit than the two-broad dimensions model (Model 5a), $\Delta \chi^2(3) = 23.76$, $p < .001$. For this reason, our further analyses only used the model with three broad dimensions (Model 5b).
The standardized parameter estimates for the broad dimensions, the general performance factor, and the uniquenesses for Model 5b can be found in Table 2. On the basis of the factor loadings, the proportion of variance accounted for by the different latent factors can be derived by calculating the mean of the squared standardized parameter estimates for each component. In our study, the general performance factor accounted for 11% of the variance in PEDRs. This is comparable to the mean for the general performance factor across the four studies by Hoffman et al. (2011). Furthermore, the broad dimension factors accounted for 5% of the variance which corresponds with the lowest value that was found in Hoffman’s et al.’s samples. Finally, the exercise factors accounted for 37% of variance in PEDRs.

The intercorrelations between the latent AC factors for Model 5b are presented in Table 3. Although the broad dimensions were significantly correlated with each other with correlations ranging from .32 to .55, these results nevertheless advocate separate albeit correlated factors.

Taken together, our results support the assumption that a model involving broad dimension factors, a general performance factor, and exercise factors is admissible and that its fit is superior to traditional models of AC ratings. Thus, the general structure suggested by Hoffman et al. (2011) was replicable for a new AC which supports the generalizability of this new model.

**Criterion-Related Validity of the AC factors**

*Academic training performance.* We then investigated the relationships between the latent AC factors of Model 5b and later academic training performance (cf. Table 3). The broad dimension factors drive \((r = .12, p < .05)\) and strategic skills \((r = .28, p < .05)\) were significantly related to academic training performance. Beyond that, the general performance factor was also significantly related to academic training performance, \(r = .20, p < .05\), as were the five exercise factors representing the short oral presentation \((r = .18, p < .05)\), the
leaderless group discussion ($r = .15, p < .05$), the short cases ($r = .13, p < .05$), the debate ($r = .14, p < .05$), and the lecture ($r = .23, p < .01$).

Next, we conducted multiple hierarchical regression analyses using the latent factor correlation matrix as input. We entered the exercise factors in the first step, the general performance factor in the second step and, finally, all broad dimensions factors in the third step (cf. Table 4).

In Step 1, the exercise factors accounted for a significant amount of criterion variance ($R^2 = .08, p < .01$). In Step 2, the general performance factor significantly improved criterion-related validity over and above the exercise factors ($\Delta R^2 = .04, p < .01$). Finally, the broad dimensions explained incremental variance over and above the exercises and the general performance factor ($\Delta R^2 = .08, p < .01$). Based on these results, the three groups of latent AC factors, namely the broad dimension factors, the general performance factor, and the exercise factors, seem to be important predictors of performance operationalized by academic training performance. Furthermore, the multiple $R$ was .44 in the final step, meaning that the different factors together explained 19% of variance in academic training performance.

Military training performance. We found that the broad dimension factor drive was significantly related to military training performance ($rs = .19, p < .01$), as was the general performance factor ($r = .25, p < .01$; cf. Table 3). Furthermore, all but one of the exercise factors had significant criterion-related validity (short oral presentation: $r = .21, p < .05$; leaderless group discussion: $r = .22, p < .01$; short cases: $r = .16, p < .05$; debate: $r = .18, p < .05$; lecture: $r = .32, p < .01$).

In Step 1 of the hierarchical regression analysis in which the exercise factors were used as predictors, all exercise factors together contributed significantly to criterion variance ($R^2 = .14, p < .01$; cf. Table 4). In Step 2, the general performance factor had incremental validity beyond the exercise factors ($\Delta R^2 = .06, p < .01$), and in the last step, the broad dimensions explained incremental variance over and above the exercise factors and the
general performance factor ($\Delta R^2 = .04, p < .01$). The multiple $R$ for all latent factors together was .49. Taken together, our results showed that all the proposed groups of factors of the new model, namely broad dimensions, the general performance factor, and the exercise factors, are criterion valid and explain incremental variance in both of our criteria.

**Nomological Network of AC factors**

We finally examined the correlations between the latent AC factors and individual difference variables in our study (Table 3). Contrary to our expectations in Hypothesis 1, no broad dimension factors were related to GMA. Furthermore, Hypothesis 2 was also not supported as neither of the broad dimension factors was correlated with the Big Five.

Hypothesis 3 was partly supported. We found strong supportive evidence for the predicted relationship between a factor indicative of leadership skills and core self-evaluations, since our factor strategic skills had moderate to high positive correlations with all core self-evaluations constructs: self-esteem: $r = .32, p < .05$, locus of control: $r = .49$, and self-efficacy: $r = .40$, both $ps < .01$. However, in contrast to our hypothesis, neither the interpersonal skills nor the drive factors were correlated with the core self-evaluations constructs.

In line with our Hypothesis 4 concerning the relationships between dimensions indicative of interpersonal skills and trait affect, the interpersonal skills factor was negatively related to negative affectivity, $r = -.20, p < .05$. In addition, the broad strategic skills factor, which was indicative of leadership skills, was moderately positively related to positive affectivity, $r = .23, p < .05$, and moderately negatively but insignificantly related to negative affectivity, $r = -.24, ns$.

As expected in Hypothesis 5, the general performance factor was significantly correlated with GMA, $r = .20, p < .01$. Furthermore, this factor was also significantly related to extraversion, $r = .17, p < .05$. 
Finally, in line with previous arguments by Lance et al. (2007) stating that exercise factors mainly represent situationally specific variance, only one significant correlation between an exercise factor and an individual difference variable was found. The motivational talk factor correlated with conscientiousness \( (r = .17, p < .01) \). However, no other exercise factors were related to individual difference variables in our study.

**Discussion**

This study makes at least three contributions to the literature. First, in a large sample of candidates from the Swiss Armed Forces, our results confirmed that the structure of AC ratings is best characterized by a model involving broad dimensions in addition to a general performance factor and exercise factors. This finding substantiates the replicability and generalizability of the recently suggested new structure of AC ratings in an AC that differed in many ways from the ACs investigated by Hoffman et al. (2011).

Second, we found that all the three different groups of latent components of this new AC structure, including the general performance factor, significantly contributed to criterion-related validity. Therefore, our findings add a caveat to previous results from Hoffman et al. (2011) that suggested that the general performance factor might not be a relevant predictor for performance criteria. Thus, this finding contributes to our understanding of the nature of the general performance factor and its relevance for predicting performance criteria. Thereby, this study also addresses calls for research concerning this factor (Kuncel & Sackett, 2014; Thornton & Rupp, 2009).

Third, we expanded the nomological network of the proposed new AC components concerning commonly studied variables like the Big Five and GMA as well as additional individual difference variables that have not been considered previously, namely the different core self-evaluations constructs and trait affectivity. Several of our assumptions concerning
the relationships between the latent AC factors and these additional individual difference variables were supported. In the following, we discuss each of our findings in detail.

**Generalizability of the Proposed AC Structure**

Our results demonstrated in an additional setting that broad dimensions and exercises are relevant components of AC PEDRs. Similar to Hoffman et al. (2011) and in line with the meta-analytic study of the internal structure of AC final dimension ratings by Meriac et al. (2014), a model with three broad dimensions, in addition to a general performance factor and exercise factors, represented the structure of our data best. Thus, our results stress that the similarity of the manifest dimensions should be taken into account when modeling AC performance and that dimension factors should not be dismissed as active components of AC performance. Moreover, a multi-faceted conceptualization involving both exercises and broad dimensions seems to be a more appropriate reflection of candidates’ AC performance. Therefore, our results suggest that AC design and interpretation should be based on information obtained from both of these components.

Similar to previous analyses of broad dimensions (Hoffman et al., 2011; Meriac et al., 2014), our model distinguished dimensions indicative of interpersonal skills from dimensions indicative of drive. A somewhat unexpected finding was that both administrative and influencing skills loaded on a common factor (strategic skills). However, this combination may be specific for our military sample. Because of the strong power hierarchy and clear obedience rules, leaders may not need strong influencing skills when they can give orders. This means that, in such a setting, influencing others may involve more cognitive or problem-solving skills and fewer relational skills.

As a caveat, it should be mentioned that the broad dimension factors accounted for a rather small percentage of variance in AC ratings in our study as compared to the general

---

1 We thank an anonymous reviewer for bringing our attention to this connection.
performance factor and the exercise factors. However, this finding is in line with results by Hoffman et al. (2011) where broad dimension factors accounted for 5% to 23% of the variance in AC ratings in the different samples that were included. Nevertheless, in our study, the broad dimension factors accounted for more variance in academic training performance than the exercise factors or the general performance factor alone. Furthermore, given our results from the nomological network of AC dimensions, our findings show that despite their relatively small contribution to variance in AC ratings the broad dimension factors are construct and criterion valid.

**Criterion-Related Validity of AC Factors**

Our findings stressed that the different latent components of AC performance contribute to criterion-related validity. Since several broad AC dimensions and most exercises explained a significant part of the variance in both of our criteria, our findings confirm that both, dimensions and exercises are important predictors of later performance that complement one another. However, it should be mentioned, that not all of the broad dimension factors and not all exercise factors were criterion valid for both criterion variables. Two broad dimension factors (drive and strategic skills) and all but one exercise factors, namely the factors representing short oral presentation, leaderless group discussion, the short cases, debate, and the lecture were significant predictors of academic training performance, while the broad dimension factor drive and all but one of the exercise factors significantly predicted military training performance. Nevertheless, in general, our findings suggest that it makes sense to use information from both sources of variance (e.g., broad dimensions and exercise factors) when interpreting AC ratings as well as for performance feedback in order to achieve valid selection decisions or developmental recommendations.

Furthermore, our study also sheds more light on the meaning of the general performance factor and its relevance as a predictor of performance. Our results showed that not only the exercise factors and the broad dimension factors (as found by Hoffman et al.,
but also the general performance factor explained a significant amount of variance in two different criteria. In addition, this factor had incremental validity beyond the exercise factors. Since the general performance factor did not significantly contribute to criterion variance in Hoffman et al. ’s study, one could have argued that this factor represents rater source effects (cf. Woehr et al., 2012) when broad dimensions are modeled. However, our findings suggest that this factor also potentially represents candidate-specific efficient and work-relevant performance throughout the AC, across all exercises or dimensions (Putka & Hoffman, 2013; Woehr et al., 2012), and that it should not generally be dismissed as a factor that may contribute to the AC’s criterion-related validity.

One possible reason for the diverging results concerning the general performance factor may lie in the composition of the sample that was tested for the present study. The candidates had very different educational backgrounds ranging from professional apprenticeships to university degrees. Thus, the overall variability in their general performance might have been larger than in previous samples, which in turn may have had an influence on the covariation of the general performance factor and training performance.

**Nomological Network**

Our results concerning the correlates of the different latent factors shed some light on the nomological network of these factors. Although several studies with final dimension ratings and with dimension composites indicated that broad dimensions should be related to the Big Five and GMA (Dilchert & Ones, 2009; Meriac et al., 2014), this was not the case in our study. However, the relationships found in these studies were low to moderate. Furthermore, a recent study by Kuncel and Sackett (2014) showed that a substantial amount of variance in dimension composites is due to a general performance factor. For this reason, it is possible that the pattern of relationships between broad dimensions, the Big Five, and GMA may change when a general performance factor is modeled.
Concerning the additional individual difference variables, all core self-evaluations constructs were moderately to highly related to the broad strategic skills factor that was indicative of leadership skills in our study. This finding likely reflects that a positive attitude towards oneself, the conviction that one can influence important outcomes, and the conviction that one is capable of doing so is important to act in a structured and purposeful manner and to be able to influence others. In their review, Judge and Kammeyer-Mueller (2011) point out that individuals who are high on core self-evaluations are more likely to set ambitious goals for themselves and are more persistent and ready to make strong efforts in their jobs. As a consequence, Judge and Kammeyer-Mueller assume that such individuals would also be effective leaders who set ambitious goals and pass on confidence to their followers. Since core self-evaluations are also related to transformational leadership (Resick, Whitman, Weingarden, & Hiller, 2009), this finding is supportive of the broad dimension factor strategic skills.

However, the finding that core self-evaluations were not related to interpersonal skills or drive was somewhat unexpected. Based on the conceptualization of the core self-evaluation constructs as beliefs about one’s self-regulatory and behavioral capacities (Johnson et al., 2008), we expected that they would be related to interpersonal skills, since a positive self-regulation is essential when interacting with others. We also expected that the core self-evaluations constructs are related to drive, since they are associated with setting motivating goals. However, possibly, in a maximum performance setting such as an AC (Ployhart, Lim, & Chan, 2001), all applicants share a high goal to succeed and are highly motivated. This could mean that the drive factor rather captures the individuals’ level of energy and not a motivational state. Furthermore, with regard to interpersonal skills, other motivational processes, for example an interaction between the core self-evaluations constructs and prosocial motivation (Grant & Wrzsniewski, 2010), might possibly be responsible for these results. Specifically, findings from Grant and Wrzsniewski suggest that individuals who are
high in core self-evaluations but who do not have a high prosocial motivation receive lower ratings than individuals who are low in core self-evaluations and have a high prosocial motivation.

Furthermore, as expected, broad dimension factors that were indicative of interpersonal behaviors (e.g., interpersonal and strategic skills) were moderately related to trait affect. Specifically, the interpersonal skills factor was negatively correlated with negative affectivity. In addition, the strategic skills factor correlated with positive affectivity and also had a moderate (but nonsignificant) negative correlation with negative affectivity. Although one of these relationships should be treated with caution due to a lack of significance, the general pattern makes sense. The results support our assumption that individuals who tend to experience more negative feelings and who tend to experience their environment in a negative way (Watson et al., 1999) would be rather less inclined to positively interact with others (Judge et al., 1998). Furthermore, our findings also suggest that individuals who are predisposed to experience positive moods and high levels of engagement (Watson et al., 1999) would be more proactive and engaged during interpersonal contacts. This assumption was especially reflected in the positive relationship between positive affectivity and the strategic skills factor that was indicative of leadership behaviors. Taken together, the relationships between trait positive affectivity and interpersonal behaviors during the AC are in the expected direction and support the construct-related validity of the broad dimension factors associated with interpersonal skills.

As expected, the general performance factor was significantly related to GMA. Furthermore, this factor was also related to extraversion. Both of these results are in line with previous meta-analytic research by Collins et al. (2003) who found high correlations between the overall assessment rating on the one hand and GMA and extraversion on the other hand. Thus, GMA was helpful in acting successfully throughout the AC. Concerning extraversion, it is likely that higher levels of extraversion were also helpful throughout all exercises of the
AC, since all exercises involved interacting with others or presenting information or ideas to others. Thus, the ability to present oneself and to approach others openly, which is associated with extraversion, could have generally led to better ratings in the AC.

The findings concerning correlations between the general performance factor and GMA and extraversion are also in line with the recent suggestion by Kuncel and Sackett (2014) that the general performance factor is related to candidates’ ability to identify criteria (ATIC; König, Melchers, Kleinmann, Richter, & Klehe, 2007). ATIC refers to the candidates’ ability to understand what is expected of them during different tasks or situations. Accordingly, it partially mediates the relationship between GMA and AC performance (Jansen et al., 2013). It is believed that, combined with the ability to act correspondingly to meet these expectations, ATIC leads to a better performance in ACs as well as on the job (Jansen et al., 2013).

Finally, our results also support previous findings concerning the exercise factors (e.g., Lance et al., 2007). As expected, the exercise factors were usually not significantly correlated with any of the individual differences constructs. Only the motivational talk factor was significantly (albeit weakly) related to conscientiousness. Even though we are hesitant to give weight to this one exception, this finding might suggest that this exercise probably had a somewhat stronger trait activation potential for this specific trait. According to Trait Activation Theory (Tett & Burnett, 2003), a situation may trigger certain trait-relevant behaviors if it provides trait-relevant signals for the individual to do so. It is plausible that a high level of conscientiousness was helpful in an exercise where one should motivate another person to do an unpleasant task. Yet, taken together, the near-absence of correlations between the exercise factors and individual differences constructs is in line with the assumption that exercise factors reflect situationally specific knowledge and performance and not individual difference constructs that are consistent across situations (Lance et al., 2007).
Practical Implications

Our study offers several practical implications for the design and administration of ACs. First, AC designers should take the similarity of narrow dimensions into account when designing and administering ACs. This means that the manifest narrow AC dimensions should be organized around empirically supported broad dimensions. Given that organizations increasingly work with competency models in various areas of human resources management, for example in personnel selection, promotion, and development (Thornton & Rupp, 2006; Campion et al., 2011), our finding concerning broad dimension factors in AC performance is also useful from the perspective of common human resource practices. Competency models are usually conceptualized as two-level hierarchies that contain subcategories of critical employee competencies (knowledge, skills, abilities, and other characteristics; KSAOs) that are arranged around broad categories (Campion et al., 2011). Our findings suggest that when using competency models for AC design and interpretation, it is reasonable from the point of view of construct-related validity of AC dimensions to rely on empirically supported broad dimension factors.

As an alternative, one might also come to the thought that narrow dimensions could be abandoned altogether in favor of broad dimensions on the basis of the present results. However, we are skeptical concerning such an idea because we believe that using narrow dimensions as data sources in ACs for the measurement of broad dimensions has several advantages. First, grouping narrow dimensions to form a common underlying broad factor does not only take the similarity of the narrow dimensions into account but also has a beneficial effect for the reliability with which the broad factors are measured. Thus, the availability of a larger number of data points (i.e., of separate ratings of the narrow dimensions) as indicators for the broad dimensions increases the reliability of these broad dimensions in comparison to when only a single PEDR for each broad dimension is collected after an exercise. Second, and in line with the first advantage, there is evidence that a high
indicator-factor ratio is beneficial for finding admissible solutions for assessment center dimensions in structural equation models (Monahan et al., 2013).

Furthermore, since exercises play a key role in predicting job-related performance, our results support previous calls to design exercises based on necessary job-relevant information to reflect the relevant primary tasks of a given position (Bank, Brock, Ramesh, & Hazucha, 2012; Jackson, 2012; Thoresen & Thoresen, 2012). Thus, by combining both sources of information, namely broad dimensions and exercises that are conceptualized as high-fidelity simulations, the predictive value of decisions based on AC results can be ensured.

Finally, decisions and (developmental) feedback to candidates should rely on both sources of information by using broad dimension scores and information concerning the candidates’ performance in specific exercises. Furthermore, candidates should also be provided with information concerning their overall performance in the AC.

**Limitations and Implications for Future Research**

Although the findings of the present study and from Hoffman et al. (2011) support the new model of the underlying structure of PEDRs from ACs, there is still a need for more rigorous tests to determine the generalizability of this model. Specifically, the particular conceptualizations of the different broad dimension factors in the present study as well as in Hoffman et al.’s samples were not conceptualized as such during the construction of the ACs but were always introduced post hoc. Thus, an even stricter test of this new conceptual structure should evaluate whether a priori specified broad dimensions can also be found in ACs that are explicitly designed according to a hierarchical competency model (cf. Campion et al., 2011). Such a test would allow additional conclusions about the practical value of the present findings with regard to their usefulness for AC construction, candidate feedback, and also for the use and alignment of human resource management practices according to broadly conceptualized higher-order competencies.
Taken together, our results advocate a mixed-model approach to AC performance and suggest that future research should involve broad dimensions when modeling candidates’ performance in the AC and on the job. Given that different aspects of situations and individuals personality probably influence human performance (Lewin, 1946), our study suggests that a multi-faceted approach may be a more appropriate reflection of candidates’ AC performance than an approach that only considers one such source (e.g., Jackson et al., 2005; Lance, 2008). Furthermore, the finding that broad dimensions are reflected in AC performance helps to align knowledge gained from managerial performance research with selection practices (Hoffman, 2012; Hofmann et al., 2011). Since managerial skills are usually conceptualized around broad dimension factors, using broad dimension factors in selection procedures may help to better predict managerial performance. Furthermore, valuable information could be gained from studying the conditions under which the different components, namely broad dimensions, exercises, and the general performance factor, better predict later performance.
References


<table>
<thead>
<tr>
<th>Model</th>
<th>admissible</th>
<th>df</th>
<th>$\chi^2$</th>
<th>SRMSR</th>
<th>RMSEA</th>
<th>CFI</th>
<th>TLI</th>
<th>$\Delta \chi^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model 1 7 dimensions</td>
<td>no$^a$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model 2 6 exercises, 7 dimensions</td>
<td>no$^b$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model 3 6 exercises</td>
<td>yes</td>
<td>284</td>
<td>578.47**</td>
<td>.036</td>
<td>.033</td>
<td>.965</td>
<td>.960</td>
<td></td>
</tr>
<tr>
<td>Model 4 6 exercises, 1 general performance factor</td>
<td>yes</td>
<td>258</td>
<td>444.27**</td>
<td>.026</td>
<td>.028</td>
<td>.978</td>
<td>.972</td>
<td>vs. Model 3: 134.20**</td>
</tr>
<tr>
<td>Model 5a 6 exercises, 2 broad dimensions, 1 general performance factor</td>
<td>yes</td>
<td>232</td>
<td>345.71**</td>
<td>.022</td>
<td>.023</td>
<td>.987</td>
<td>.981</td>
<td>vs. Model 4: 98.56**</td>
</tr>
<tr>
<td>Model 5b 6 exercises, 3 broad dimensions, 1 general performance factor</td>
<td>yes</td>
<td>229</td>
<td>321.95**</td>
<td>.023</td>
<td>.021</td>
<td>.989</td>
<td>.984</td>
<td>vs. Model 5a: 23.76**</td>
</tr>
</tbody>
</table>

*Note.* SRMSR = standardized root mean squared residual; RMSEA = root mean squared error of approximation; CFI = comparative fit index; TLI = Tucker-Lewis Index.

$^a$ model not admissible due to a not positive definite covariance matrix. $^b$ model unidentified.

** $p < .001$
### Table 2

**Standardized Parameter Estimates for Model 5b**

<table>
<thead>
<tr>
<th>PEDR</th>
<th>Exercise factors</th>
<th>Broad dimensions</th>
<th>GP</th>
<th>Unique</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>SOP</td>
<td>LGD</td>
<td>MOT</td>
<td>SHC</td>
</tr>
<tr>
<td>SOP Achievement motivation</td>
<td>.60**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SOP Analysis and planning</td>
<td>.53**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SOP Oral communication</td>
<td>.39**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LGD Personal attitude</td>
<td>.67**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LGD Social skills</td>
<td>.62**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LGD Oral communication</td>
<td>.53**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LGD Influencing others</td>
<td>.72**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MOT Personal attitude</td>
<td>.65**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MOT Achievement motivation</td>
<td>.51**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MOT Social skills</td>
<td>.61**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MOT Oral communication</td>
<td>.45**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MOT Influencing others</td>
<td>.64**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SHC Personal attitude</td>
<td>.61**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SHC Achievement motivation</td>
<td>.56**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SHC Analysis and planning</td>
<td>.69**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SHC Social skills</td>
<td>.46**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SHC Influencing others</td>
<td>.70**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DEB Personal attitude</td>
<td>.72**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DEB Social skills</td>
<td>.49**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DEB Oral communication</td>
<td>.53**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DEB Influencing others</td>
<td>.68**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PRE Achievement motivation</td>
<td>.64**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PRE Analysis and planning</td>
<td>.76**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PRE Oral communication</td>
<td>.54**</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Note. SOP = short oral presentation; LGD = leaderless group discussion; MOT = motivational talk; SHC = short cases; DEB = debate; PRE = lecture on a topic of military pedagogy; INT = interpersonal skills; DRI = drive; SS = strategic skills; GP = general performance factor.

* p < .05, ** p < .01.
### Table 3

**Correlations between the latent factors of Model 5b and all study variables**

<table>
<thead>
<tr>
<th>Variables</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
<th>16</th>
<th>17</th>
<th>18</th>
<th>19</th>
<th>20</th>
<th>21</th>
<th>22</th>
<th>23</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Short presentation</td>
<td>(-)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. LGD</td>
<td>.34**</td>
<td>(-)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Motivational talk</td>
<td>.22*</td>
<td>.29**</td>
<td>(-)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Short cases</td>
<td>.13</td>
<td>.28**</td>
<td>.19**</td>
<td>(-)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Debate</td>
<td>.20*</td>
<td>.37**</td>
<td>.28**</td>
<td>.22**</td>
<td>(-)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Lecture</td>
<td>.24**</td>
<td>.25**</td>
<td>.14*</td>
<td>.26**</td>
<td>.25**</td>
<td>(-)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. Interpersonal skills</td>
<td>.00</td>
<td>.00</td>
<td>.00</td>
<td>.00</td>
<td>.00</td>
<td>.00</td>
<td>.00</td>
<td>.00</td>
<td>.44</td>
<td>(-)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. Drive</td>
<td>.00</td>
<td>.00</td>
<td>.00</td>
<td>.00</td>
<td>.00</td>
<td>.00</td>
<td>.00</td>
<td>.32</td>
<td>.55</td>
<td>(-)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9. Strategic skills</td>
<td>.00</td>
<td>.00</td>
<td>.00</td>
<td>.00</td>
<td>.00</td>
<td>.00</td>
<td>.00</td>
<td>.00</td>
<td>.00</td>
<td>.00</td>
<td>(-)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10. GP</td>
<td>.00</td>
<td>.00</td>
<td>.00</td>
<td>.00</td>
<td>.00</td>
<td>.00</td>
<td>.00</td>
<td>.00</td>
<td>.20**</td>
<td>(.79)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11. GMA&lt;sup&gt;a&lt;/sup&gt;</td>
<td>.05</td>
<td>-.01</td>
<td>.01</td>
<td>.06</td>
<td>.03</td>
<td>.08</td>
<td>.06</td>
<td>.08</td>
<td>.00</td>
<td>.08*</td>
<td>.16**</td>
<td>(.60)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12. Extraversion&lt;sup&gt;b&lt;/sup&gt;</td>
<td>-.13</td>
<td>.04</td>
<td>-.06</td>
<td>-.03</td>
<td>-.04</td>
<td>-.09</td>
<td>.09</td>
<td>.09</td>
<td>.19</td>
<td>.17*</td>
<td>-.01</td>
<td>(.74)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13. Emot. Stability&lt;sup&gt;b&lt;/sup&gt;</td>
<td>.11</td>
<td>.06</td>
<td>.06</td>
<td>.08</td>
<td>-.03</td>
<td>.00</td>
<td>.06</td>
<td>.08</td>
<td>.00</td>
<td>.08*</td>
<td>.16**</td>
<td>(.60)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14. Openness&lt;sup&gt;b&lt;/sup&gt;</td>
<td>.10</td>
<td>.09</td>
<td>.01</td>
<td>.01</td>
<td>.11</td>
<td>-.01</td>
<td>-.04</td>
<td>-.05</td>
<td>-.07</td>
<td>.02</td>
<td>-.15**</td>
<td>.28**</td>
<td>.06</td>
<td>(.75)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15. Agreeableness&lt;sup&gt;b&lt;/sup&gt;</td>
<td>.13</td>
<td>.13</td>
<td>.02</td>
<td>.15</td>
<td>.13</td>
<td>.08</td>
<td>.05</td>
<td>-.01</td>
<td>-.15</td>
<td>-.16</td>
<td>.05</td>
<td>-.04</td>
<td>.05</td>
<td>(.54)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16. Conscientiousness&lt;sup&gt;b&lt;/sup&gt;</td>
<td>.11</td>
<td>-.06</td>
<td>.17*</td>
<td>.06</td>
<td>.10</td>
<td>.12</td>
<td>-.07</td>
<td>.00</td>
<td>-.07</td>
<td>-.16**</td>
<td>.01</td>
<td>.14**</td>
<td>.04</td>
<td>-.01</td>
<td>(.77)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17. Self esteem&lt;sup&gt;c&lt;/sup&gt;</td>
<td>.09</td>
<td>.01</td>
<td>-.03</td>
<td>.08</td>
<td>.06</td>
<td>.03</td>
<td>-.01</td>
<td>-.04</td>
<td>.32*</td>
<td>-.01</td>
<td>.02</td>
<td>.17**</td>
<td>.49**</td>
<td>.11*</td>
<td>.05</td>
<td>.25**</td>
<td>(.67)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18. Locus of control&lt;sup&gt;d&lt;/sup&gt;</td>
<td>-.02</td>
<td>-.08</td>
<td>-.02</td>
<td>-.09</td>
<td>.02</td>
<td>-.03</td>
<td>.12</td>
<td>.01</td>
<td>.49**</td>
<td>-.13</td>
<td>.00</td>
<td>.08</td>
<td>.23**</td>
<td>.07</td>
<td>.09</td>
<td>.22**</td>
<td>.32**</td>
<td>(.61)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>19. Self efficacy&lt;sup&gt;e&lt;/sup&gt;</td>
<td>.08</td>
<td>-.11</td>
<td>.02</td>
<td>.03</td>
<td>.08</td>
<td>.00</td>
<td>-.06</td>
<td>-.09</td>
<td>.40**</td>
<td>.02</td>
<td>.02</td>
<td>.07</td>
<td>.33**</td>
<td>.17**</td>
<td>.02</td>
<td>.27**</td>
<td>.41**</td>
<td>.32**</td>
<td>(.69)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20. PA&lt;sup&gt;f&lt;/sup&gt;</td>
<td>-.06</td>
<td>.04</td>
<td>-.05</td>
<td>.02</td>
<td>.00</td>
<td>.10</td>
<td>-.04</td>
<td>.23*</td>
<td>.07</td>
<td>-.07</td>
<td>-.32**</td>
<td>.40**</td>
<td>.26**</td>
<td>-.01</td>
<td>.44**</td>
<td>.44**</td>
<td>.29**</td>
<td>.38**</td>
<td>(.79)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>21. NA&lt;sup&gt;f&lt;/sup&gt;</td>
<td>-.08</td>
<td>.00</td>
<td>.09</td>
<td>.01</td>
<td>.06</td>
<td>-.04</td>
<td>-.20*</td>
<td>.03</td>
<td>-.23</td>
<td>-.02</td>
<td>-.09</td>
<td>-.03**</td>
<td>-.46*</td>
<td>-.01</td>
<td>-.30**</td>
<td>-.12*</td>
<td>-.30**</td>
<td>-.16**</td>
<td>-.26**</td>
<td>-.16**</td>
<td>(.66)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>22. Academic training performance&lt;sup&gt;d&lt;/sup&gt;</td>
<td>.18*</td>
<td>.15*</td>
<td>.12</td>
<td>.13*</td>
<td>.14*</td>
<td>.23**</td>
<td>.08</td>
<td>.12*</td>
<td>.28*</td>
<td>.20*</td>
<td>.22**</td>
<td>-.11</td>
<td>.01</td>
<td>-.02</td>
<td>.05</td>
<td>.03</td>
<td>.16</td>
<td>.09</td>
<td>.07</td>
<td>.02</td>
<td>-.13</td>
<td>(-)</td>
<td></td>
</tr>
<tr>
<td>23. Military training performance&lt;sup&gt;d&lt;/sup&gt;</td>
<td>.21*</td>
<td>.22**</td>
<td>.10</td>
<td>.16*</td>
<td>.18*</td>
<td>.32**</td>
<td>.08</td>
<td>.19**</td>
<td>.09</td>
<td>.25**</td>
<td>.21**</td>
<td>-.08</td>
<td>.13*</td>
<td>-.01</td>
<td>-.08</td>
<td>.06</td>
<td>.21*</td>
<td>.05</td>
<td>.13</td>
<td>-.02</td>
<td>-.08</td>
<td>.52**</td>
<td>(-)</td>
</tr>
</tbody>
</table>

**Note.** Values in the diagonal show internal consistencies. LGD = leaderless group discussion; GP = general performance factor; PA = positive affectivity; NA = negative affectivity.

<sup>a</sup>n = 734, <sup>b</sup>n = 704, <sup>c</sup>n = 432, <sup>d</sup>n = 430, <sup>e</sup>n = 431, <sup>f</sup>n = 429, <sup>g</sup>n = 500, z-standardized. <sup>h</sup>n = 514.

* p < .05, ** p < .01.
Table 4

*Academic and military training performance regressed on the latent AC factors*

<table>
<thead>
<tr>
<th></th>
<th>( \beta )</th>
<th>( R^2 )</th>
<th>( \Delta R^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Academic training performance (N = 500)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Step 1</td>
<td></td>
<td>( .08^{**} )</td>
<td>( .08^{**} )</td>
</tr>
<tr>
<td>Short oral presentation</td>
<td>.10*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leaderless group discussion</td>
<td>.02</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Motivational talk</td>
<td>.05</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Debate</td>
<td>.05</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Short cases</td>
<td>.05</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lecture</td>
<td>.17^{**}</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Step 2</td>
<td></td>
<td>( .12^{**} )</td>
<td>( .04^{**} )</td>
</tr>
<tr>
<td>General performance</td>
<td>.20^{**}</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Step 3</td>
<td></td>
<td>( .20^{**} )</td>
<td>( .08^{**} )</td>
</tr>
<tr>
<td>Interpersonal skills</td>
<td>-.01</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Drive</td>
<td>-.05</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Strategic skills</td>
<td>.31^{**}</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Military training performance (N = 514)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Step 1</td>
<td></td>
<td>( .14^{**} )</td>
<td>( .14^{**} )</td>
</tr>
<tr>
<td>Short oral presentation</td>
<td>.13</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leaderless group discussion</td>
<td>.10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Motivational talk</td>
<td>-.01</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Debate</td>
<td>.08</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Short cases</td>
<td>.07</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lecture</td>
<td>.13^{**}</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Step 2</td>
<td></td>
<td>( .20^{**} )</td>
<td>( .06^{**} )</td>
</tr>
<tr>
<td>General performance</td>
<td>.22^{**}</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Step 3</td>
<td></td>
<td>( .24^{**} )</td>
<td>( .04^{**} )</td>
</tr>
<tr>
<td>Interpersonal skills</td>
<td>-.02</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Drive</td>
<td>.21^{**}</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Strategic skills</td>
<td>-.01</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* *p < .05; ** p < .01.*