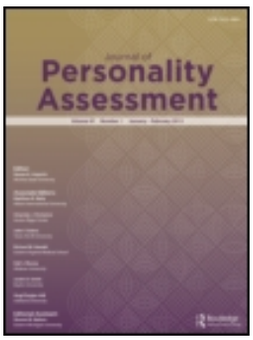


Routledge
Taylor & Francis Group

Journal of Personality Assessment

ISSN: 0022-3891 (Print) 1532-7752 (Online) Journal homepage: http://www.tandfonline.com/loi/hjpa20

Factors Underlying the Psychological and Behavioral Characteristics of Office of Strategic Services Candidates: The *Assessment of Men* Data Revisited

Mark F. Lenzenweger

To cite this article: Mark F. Lenzenweger (2015) Factors Underlying the Psychological and Behavioral Characteristics of Office of Strategic Services Candidates: The *Assessment of Men* Data Revisited, Journal of Personality Assessment, 97:1, 100-110, DOI: 10.1080/00223891.2014.935980

To link to this article: https://doi.org/10.1080/00223891.2014.935980

Published online: 18 Jul 2014.

Submit your article to this journal

Article views: 177

View Crossmark data

Citing articles: 1 View citing articles

Full Terms & Conditions of access and use can be found at
http://www.tandfonline.com/action/journalInformation?journalCode=hjpa20

Journal of Personality Assessment, *97*(1), 100–110, 2015
Copyright © Taylor & Francis Group, LLC
ISSN: 0022-3891 print / 1532-7752 online
DOI: 10.1080/00223891.2014.935980

Routledge
Taylor & Francis Group

Factors Underlying the Psychological and Behavioral Characteristics of Office of Strategic Services Candidates: The *Assessment of Men* Data Revisited

MARK F. LENZENWEGER[1,2]

[1]*Department of Psychology, The State University of New York at Binghamton*
[2]*Department of Psychiatry, Weill Cornell Medical College, Cornell University*

During World War II, the Office of Strategic Services (OSS), the forerunner of the Central Intelligence Agency, sought the assistance of clinical psychologists and psychiatrists to establish an assessment program for evaluating candidates for the OSS. The assessment team developed a novel and rigorous program to evaluate OSS candidates. It is described in *Assessment of Men: Selection of Personnel for the Office of Strategic Services* (OSS Assessment Staff, 1948). This study examines the sole remaining multivariate data matrix that includes all final ratings for a group of candidates ($n = 133$) assessed near the end of the assessment program. It applies the modern statistical methods of both exploratory and confirmatory factor analysis to this rich and highly unique data set. An exploratory factor analysis solution suggested 3 factors underlie the OSS assessment staff ratings. Confirmatory factor analysis results of multiple plausible substantive models reveal that a *3-factor model* provides the best fit to these data. The 3 factors are emotional/interpersonal factors (social relations, emotional stability, security), intelligence processing (effective IQ, propaganda skills, observing and reporting), and agency/surgency (motivation, energy and initiative, leadership, physical ability). These factors are discussed in terms of their potential utility for personnel selection within the intelligence community.

During the early stage of World War II, the United States formed, via executive order from President Franklin D. Roosevelt in 1941, the Office of the Coordinator of Information, which then became the Office of Strategic Services (OSS) under the direction of General William J. Donovan (Waller, 2011). One purpose of the OSS was to function as the intelligence service for the United States for acquisition and analysis of intelligence on matters of national security concern to the country. As a second purpose, the OSS was authorized to carry out clandestine special operations, including subversion, propaganda, and psychological warfare operations, behind enemy lines in Europe and the Far East, in an effort to advance the war effort against Nazi Germany and Japan. Thus, the OSS served both intelligence and operational functions. The OSS was dissolved almost immediately after the end of World War II. The intelligence and some operational functions of the OSS were taken over largely in 1947 by the newly formed Central Intelligence Agency (CIA). The special operations component of the OSS is recognized as the progenitor of the U.S Army Special Forces.

The very first OSS officers were placed in the field not long after the formation of the organization, and the selection of these early OSS officers was understandably not particularly systematic. The organization was entirely new; it was tasked with goals and objectives that developed rapidly, and time pressures were immense. Initial reports back from the field regarding poor performance by some OSS officers (i.e., due to incompetence or psychological dysfunction in high-threat/stress situations) suggested the need for more thorough and detailed assessment of OSS candidates to improve the selection process before deployment on assignment (Banks, 1995; MacKinnon, 1974/1980; OSS Assessment Staff, 1948). Given that the United States did not have a systematically developed, centralized institutional commitment to an intelligence agency prior to the OSS, there were no explicit psychological procedures or guidelines for use in personnel selection for members of such an agency. This situation, however, was not unique to the OSS. It is worth noting that psychological selection criteria for use within the U.S. military, generally, were not well developed either. For example, the psychological health of armed forces inductees was an area where psychological screening for mental health saw development of only rudimentary technologies and limited application of them in both World War I (e.g., Woodworth Personal Data Sheet; see Strecker, 1944; Woodworth, 1919) and World War II (e.g., the Neuropsychiatric Screening Adjunct; see Pols & Oak, 2007, for review). The Personal Data Sheet was developed relatively late in World War I and did not see extensive use. The Neuropsychiatric Screening Adjunct saw minimal application during World War II and was not viewed favorably as an effective screening tool (Banks, 1995). Clearly, selection of personnel suitable for intelligence work—either operational or analytic—was not a focus of applied psychological branches either within or outside the military. In short, given the feedback from the field regarding the performance of some (but, by no means all) early OSS officers, the OSS found itself in a relatively novel position with the need for help with assessment and selection in a circumstance where time was of the essence. Professional assistance in the assessment and evaluation of potential OSS candidates began in earnest in 1943 (MacKinnon, 1974/1980; OSS Assessment Staff, 1948).

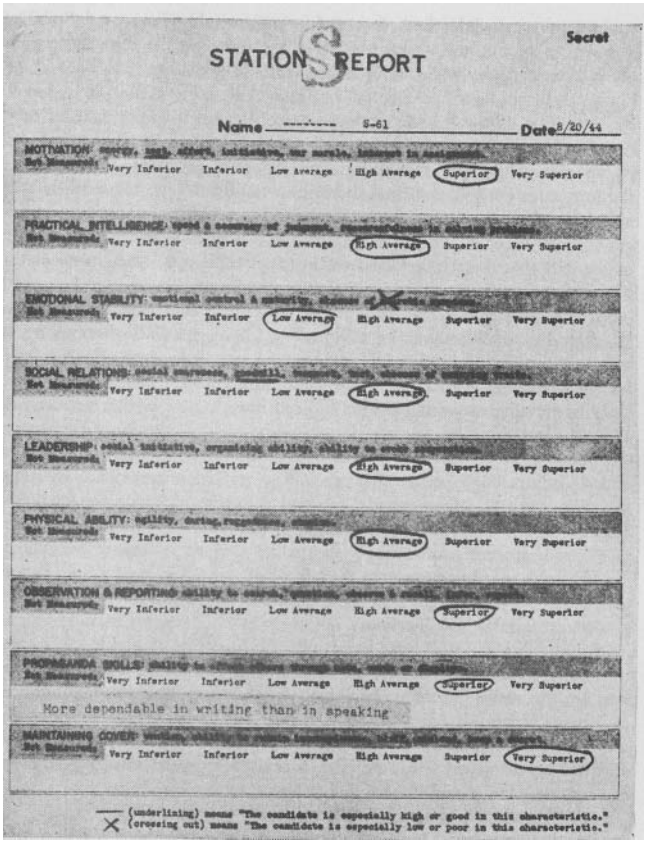
Received December 5, 2013; Revised May 5, 2014.

Address correspondence to Mark F. Lenzenweger, Department of Psychology, The State University of New York at Binghamton, Science IV (G-08), Binghamton, NY 13902-6000; Email: mlenzen@binghamton.edu

In response to the need for professional assistance in developing a more thorough psychological and behavioral assessment adjunct to selection, the OSS reached out to a number of prominent clinical psychologists and psychiatrists in the academic community within the United States (MacKinnon, 1974/1980; OSS Assessment Staff, 1948). This established the connection between clinical psychology, with its focus on developing assessment approaches, and the nascent U.S. intelligence community. There had been no comparable prior instance (prior to World War II) in the fields of clinical psychology, personnel psychology, or clinical psychiatry fields where intensive study of individuals was carried out for the stated purpose of selection for likely suitability as an intelligence officer or special operations personnel (Banks, 2006; Williams, Picano, Roland, & Bartone, 2012; see also Butcher, 2010). In contrast to this situation in the United States, both British and German psychologists and psychiatrists had been active in assisting in the selection of officers for the military during World War I and prior to World War II (Banks, 1995). In fact, the OSS received input in 1943 regarding the nature of the British War Office Selection Boards, and this served as an impetus for the OSS to engage with U.S. psychologists and psychiatrists (MacKinnon, 1974/1980).

Many well-known psychologists and psychiatrists (see Handler, 2001; OSS Assessment Staff, 1948) participated in the development of the assessment protocol that was used in the OSS assessment program. The lead was taken by Henry A. Murray, the Harvard psychologist whose personality theory and assessment approach had gained considerable traction in clinical psychology and personality at the time (Murray, 1938). It is important to bear in mind the scientific context in which the OSS assessment team began its work. At that time, classical Freudian "drive theory" psychoanalysis (not its amended, extended, or derivative versions to appear in ensuing decades) held sway in the minds of most psychologists and psychiatrists, operant (Skinnerian) behaviorism was in its infancy and contributed little to human clinical assessment, and there were virtually no factorial models of normal personality such as we have today (e.g., the popular three-factor, five-factor, and seven-factor variants). The challenge set for the OSS assessment team, which set up initial operations at the Willard Estate in Fairfax, VA, just outside of Washington, DC, at what was called Station S (for Schools and Training), was to define the criteria that would be central to the selection of OSS personnel. In short, the question faced by the assessors was this: What should be the bases for selection of a potentially successful OSS officer? What psychological or behavioral features should a successful OSS officer possess or reveal on assessment? The particularly difficult aspect of this challenge was that the fundamental nature of the ultimate "to be predicted" criterion (i.e., the "successful" OSS officer) was not known at the time the assessment enterprise was undertaken, nor could it have been known with certainty in advance (OSS Assessment Staff, 1948). In short, the OSS assessment staff did not know precisely what they were asked to predict in developing selection criteria for candidates. Nonetheless, the clinical psychology and psychiatry staff brought on to establish the assessment approach designed an intensive evaluation program that could be useful in tapping into many aspects of psychological, behavioral, and social functioning they thought to be important in selecting OSS officers. The assessment approach they relied on would become known eventually as the "assessment center" method (MacKinnon, 1974/1980), wherein candidates were evaluated intensively on many behavioral tasks, psychological tests, and stressful situations over several days. It was thought that these tasks, tests, and situations would mirror, in theory, circumstances or conditions in which actual OSS officers would be required to function in the field. The OSS assessment staff sought to evaluate the "whole person" in their assessment and selection protocol—cognitive, emotional, personality, and behavioral characteristics—modeled in part on Murray's (1938) conceptualization of the human personality.

The assessment protocol at Station S has been described in great detail elsewhere (Handler, 2001; OSS Assessment Staff, 1948) and is not repeated here. In brief, to provide some context, small groups of candidates (usually 18) were transported twice weekly from Washington, DC, to Station S after being instructed to assume new identities and to not reveal anything about their actual identities while at Station S. The candidates were told they would be given tests and be asked questions by psychologists and psychiatrists over a period of 3 days. Once at Station S, the candidates were assessed on 85 to 90 assessments, which were reduced conceptually and used to inform the final ratings of the candidates on 10 core dimensions. Some behavioral tasks were especially interesting, such as the "construction situation," which involved a candidate attempting to build a small structure within 10 min in collaboration with two others, who were actually OSS assessment staff "stooges." The job of the "stooges" in this construction task was to be as maximally frustrating and difficult as they could be in relation to the candidate who was actually trying to build the small structure. This task, as well as others, yielded rich observational data about the candidates for the assessment team (see OSS Assessment Staff, 1948, for detail). Overall, the assessment staff gathered their many observations, test results, and interviews into seven broad conceptual psychological categories, which were rated: (a) motivation for the assignment, (b) energy and initiative, (c) practical (effective) intelligence, (d) emotional stability, (e) social relationships, (f) leadership, and (g) security (i.e., ability to keep secrets, ability to bluff, maintain cover). Three other characteristics that were a focus were rated as well: (h) physical ability, (i) observing and reporting, and (j) propaganda skills (see OSS Assessment Staff, 1948, for greater detail). Ratings for each candidate on these 10 categories (using a 6-point scale ranging from *very inferior* to *very superior*) were made by the assessment staff at Station S in what was termed the staff conference. The staff conference was described by the OSS Assessment Staff (1948) as the critical event in the evaluation protocol. A summary report was generated for each candidate (see Figure 1) that provided the final clinical judgments made in the staff conference. The goals of the staff conference were to integrate and synthesize all that had been learned about the totality of a candidate's personality and to render a prediction regarding the possible future performance of the candidate in an OSS role. The staff conferences were characterized by "vigor and vitality," and "free expression," and were "often heated" (OSS Assessment Staff, 1948, pp. 217–218). There were more than 50 professional staff at Station


Secret

STATION S REPORT

Name _____ S-61 Date 8/20/44

MOTIVATION: energy, drive, ardour, initiative, ear marks, interest in assignment.
Not Measured: Very Inferior Inferior Low Average High Average (Superior) Very Superior

PRACTICAL INTELLIGENCE: speed & accuracy of judgment, resourcefulness in solving problems.
Not Measured: Very Inferior Inferior Low Average (High Average) Superior Very Superior

EMOTIONAL STABILITY: emotional control & maturity, absence of [illegible]
Not Measured: Very Inferior Inferior (Low Average) High Average Superior Very Superior

SOCIAL RELATIONS: social awareness, sociality, tact, skill in [illegible]
Not Measured: Very Inferior Inferior Low Average (High Average) Superior Very Superior

LEADERSHIP: social initiative, organizing ability, ability to evoke cooperation.
Not Measured: Very Inferior Inferior Low Average (High Average) Superior Very Superior

PHYSICAL ABILITY: agility, daring, [illegible]
Not Measured: Very Inferior Inferior Low Average (High Average) Superior Very Superior

OBSERVATION & REPORTING: ability to observe, question, observe & report.
Not Measured: Very Inferior Inferior Low Average High Average (Superior) Very Superior

PROPAGANDA SKILLS: ability to [illegible]
Not Measured: Very Inferior Inferior Low Average High Average (Superior) Very Superior

More dependable in writing than in speaking

MAINTAINING COVER: [illegible]
Not Measured: Very Inferior Inferior Low Average High Average Superior (Very Superior)

—— (underlining) means "The candidate is especially high or good in this characteristic."
X (crossing out) means "The candidate is especially low or poor in this characteristic."

FIGURE 1.—Summary sheet from Station S assessment used to provide final Staff Conference ratings for candidates. (Note supplementary comment for Propaganda Skills, "More dependable in writing than in speaking" for this candidate). On this version of the Station S Report form, "motivation" is rated as a cluster that included "energy and initiative," although "motivation" and "energy and initiative" were typically distinguished and treated as separate variables at Station S. They were treated as separate variables in all analyses in this report. (Resolution of image somewhat degraded in the original presented in *Assessment of Men* [Office of Strategic Services Assessment Staff, 1948]). The original publication is now classified as a public domain document.

S, the main assessment center (there were several others), and more than 5,000 candidates were assessed there in a period of about 20 months across all OSS assessment sites (Handler, 2001).[1]

The OSS assessment program at Station S and the other assessment stations clearly produced an enormous amount of empirical psychological data. However, most of these data were never analyzed fully in any detailed manner. Rudimentary analyses were presented in the Appendices of the *Assessment of Men* (OSS Assessment Staff, 1948), with some secondary analyses examining performance outcomes (available after the war) being conducted some years later by personality psychologist Jerry Wiggins (1973). The multivariate array of the psychological data generated by the OSS assessments, which likely harbors interesting factors, was subject to a limited multivariate analysis right after the war, namely an exploratory factor analysis (EFA). That factor analysis sought to reduce the large number of assessment variables to a smaller set of underlying factors. This early analysis, carried out by the OSS Assessment team and reported in *Assessment of Men* (OSS Assessment Staff, 1948), was done essentially by hand and, as such, was not as richly or precisely conducted as can be done with modern methods. The factor analysis solution reported in the *Assessment of Men* was entirely exploratory in nature. Moreover, the methods at the time did not allow the original psychological investigators to determine which of several competing substantive models might provide the best fit to their data. This was so because computational technology was limited in the late 1940s, which limited the form of EFAs that could be done. Also, importantly, the statistical approach known as confirmatory factor analysis (CFA) simply did not exist at that time.

This study, therefore, seeks to apply modern statistical methods to the analysis of the multivariate data available from the original OSS assessments done at Station S for a subset of candidates evaluated there. The data available for this study are contained in a correlation matrix relating the variables generated by the staff conference results for the candidates as published in *Assessment of Men* (OSS Assessment Staff, 1948). Clearly, the assessments done at Station S were remarkable for their richness, and they generated extraordinarily unique data. It makes scientific as well as practical sense to use modern factor analytic methods to extract as much meaning and direction from those data as possible. It would be useful to understand the nature of the latent structure underlying the final characterizations of the OSS candidates. Moreover, from the standpoint of personnel selection, a more thorough understanding of those factors involved in the selection process might prove useful to parties with an interest in personnel selection in the intelligence community. What makes this particular correlation matrix worthy of detailed attention is that although many individuals passed through the assessment center at Station S (and other locations), it is reported that all records pertaining to the candidate evaluations conducted at Station S were ordered destroyed at the end of the war (Handler, 2001). Thus, this matrix is perhaps the only multivariate vestige of the herculean effort embodied in this important chapter of psychological assessment, an effort that gave birth to the modern assessment center approach that is in use today (Bhyham, 2002). The insights gleaned from the OSS program are still deemed important, not only from a historical perspective, but because they continue to inform operational personnel selection today (Banks, 2006). This study seeks to explore the original OSS data further with the intention of providing greater quantitative and conceptual clarity on the general psychological processes tapped through the OSS assessment program and, in doing so, make a contribution to modern personnel evaluation and selection discussions in the intelligence and special operations communities.

[1]Not all of those who served in the OSS were processed through the assessment protocol described here even after it was established at Station S and other locations. Some members were recruited directly into the OSS from within active combat zones without completing an assessment. Such recruits needed only to secure a release from their previously assigned unit to move into the OSS (John Behling, PhD, personal communication, January 30, 2013).

METHOD

Subjects

The subjects for this study were 133 candidates from several of the final OSS candidate classes evaluated at Station S. The data from these 133 subjects were presented in the form of a published correlation matrix in the Appendix (OSS Assessment Staff, 1948, Appendix B, p. 510). The matrix is based on complete data on all 133 subjects for all measures.

Data Structure

The battery of measures, assessment devices, behavioral task situations, and interviews (including stress interviews) used to assess OSS candidates was extensive and described in extensive detail in *Assessment of Men* (OSS Assessment Staff, 1948). This massive corpus of assessment data was used as the basis for the clinical ratings on 11 core dimensions for each candidate by the original assessment staff as described earlier. The data analyzed for this study is a correlation matrix (Table 1) relating these 11 variables (with unities [1.00] placed initially in the diagonal). The original correlation matrix contained 11 psychological, personality, or behavioral variables (the 10 variables plus an "overall" rating; see later). The dimensions are: (a) motivation for the assignment,[2] (b) energy and initiative, (c) practical (effective) intelligence, (d) emotional stability, (e) social relationships, (f) leadership, (g) security (i.e., ablilty to keep secrets, ability to bluff, maintain cover), (h) physical ability, (i) observing and reporting, and (j) propaganda skills. Each of these corresponded to a quantitative dimension on which the candidates were evaluated, and summary ratings were made by the OSS assessment staff. An important concern regarding the EFA done originally by the OSS assessment staff was that it used the correlation matrix described here, but it also included an additional (11th) variable. The additional variable was described as an "overall" summary rating. This overall rating was clearly highly redundant with the other variables reported because it was statistically infused with the ratings of the other 10 variables under consideration. Inclusion of the overall rating variable in the original matrix probably introduced a statistical artifact into the original analysis, which probably interfered with model estimation and could have contributed to error inflation in the factor analysis. Moreover, perhaps more important, the exact meaning of the overall rating was clearly elusive even to the OSS assessment staff, who described it as "an estimate of the total potentialities of the candidates for meeting the challenges of life—exceedingly vague and difficult concept to define" (OSS Assessment Staff, 1948, p. 217). For the purposes of this analysis, the overall variable has been eliminated. The subject to variable ratio for this study is therefore 13.3 to 1, which is acceptable, and the sample size of 133 is within the bounds of sample size that allow for good factor recovery (when fewer factors are retained; MacCallum, Widaman, Zhang, & Hong, 1999; Preacher & MacCallum, 2002).

Statistical Analysis

The correlation matrix in Table 1 was statistically reanalyzed in a two-step process. The first step of the analysis consisted of an EFA (principal axis factoring), followed by both orthogonal (Varimax) and oblique rotations. The Kaiser criterion (eigenvalue > 1.0) was used as the threshold for factor retention. This analysis was conducted to determine the latent structure of the data in a manner comparable to that reported originally in *Assessment of Men,* but using modern computational methods not available to the original investigators (and removal of the "overall" rating). The original EFA reported in *Assessment of Men* is only vaguely described with respect to technical detail (e.g., extraction method was likely the "centroid" method; rotation method was not specified (if used); values used in the matrix diagonal were not specified). The centroid method used in that analysis was intended only to approximate a factor analytic solution as the centroid method represented something of a "computational compromise" (Harman, 1976, p. 166; that is, it could be done easily on paper, but at the expense of technical precision. With the development of fast and efficient computer technology along with abundant storage space, the centroid method of factor analysis was replaced by the more precise method of principal axis factoring (the EFA technique used in this study) and related methods. Moreover, the centroid method, done manually, required extraordinary vigilance owing to computing complexities (e.g., numerous sign reversals on many values across many computational steps) and, as such, it was an approach well known to invite human error. Thus, it seems prudent to conduct a new EFA on the original data using more precise modern methods that are more reliable than the defunct centroid method.

The second step in the data analysis used CFA to estimate a series of nested models of the latent structures plausibly underlying the observed correlations. Unlike the EFA approach, which is not driven by a priori specified substantive models of how the observed data are structured, a CFA approach specifically requires the investigator to specify models prior to the analysis. In short, CFA requires one to articulate how the observed covariances (or correlations) among the variables can be driven by underlying factors, or latent structures, and such models should have a substantive or theoretical basis. The specified theoretical model is evaluated against the observed data in terms of how well the model fits the observed data fit. A theoretical model that produces a relatively close fit to the data is sought, and the degree of fit is evaluated quantitatively. The factor solution obtained with CFA estimation is both unique and direct and, therefore, no rotation of the solution is necessary to interpret it theoretically. When several theoretical models are considered plausible, each of the models is estimated separately in CFA, and the results of these analyses

[2] "Motivation for assignment" was thought to consist of two distinct variables throughout most of the OSS assessment program's duration; namely (a) desire to accomplish an assignment in the OSS, and (b) level of energy and initiative in relation to achieving goals (see OSS Assessment Staff, 1948, p. 233). At some points in the program, the two variables were treated as a single composite for some candidate classes. This can be seen, for example, in the rating sheet displayed in Figure 1, which contains only a "motivation" category, which merged the two variables. However, for most of the time, motivation for assignment and energy and initiative were assessed and considered separately at Station S. Importantly, the data used in this study were derived from a period in the OSS assessment program where these two variables were unambiguously separated. This allows for motivation for assignment and energy and initiative to be treated as two distinct variables in the statistical analyses in this report.

TABLE 1.—Intercorrelations among the final assessment variables in the Office of Strategic Services assessment program at Station S.

	Motivation (1)	Energy & Initiative (2)	Effective Intelligence (3)	Emotional Stability (4)	Social Relations (5)	Leadership (6)	Physical Ability (7)	Security (8)	Observing & Reporting (9)	Propaganda Skills (10)
1	—									
2	.47	—								
3	.31	.56	—							
4	.43	.53	.23	—						
5	.39	.38	.27	.62	—					
6	.44	.72	.65	.48	.44	—				
7	.26	.41	.06	.34	.38	.21	—			
8	.27	.21	.11	.37	.33	.16	.13	—		
9	.23	.31	.63	.22	.32	.32	.07	.18	—	
10	.37	.36	.70	.21	.28	.51	–.07	.21	.53	—

Note. $n = 133$. The values are Pearson product–moment correlation coefficients as reported in Table 53 (p. 511) in *Assessment of Men: Selection of Personnel for the Office of Strategic Services* (OSS Assessment Staff, 1948). All 133 men were rated on all variables.

are then statistically contrasted to determine which of those models tested fits the data best (see Lenzenweger, Dworkin, & Wethington, 1989, for an extensive discussion of the merits of CFA over EFA approaches). The CFA approach employed here made use of maximum-likelihood estimation for computations and several well-established indexes or procedures for the evaluation of the fit between stand-alone theoretical models and the observed results (goodness of fit chi-square, Akaike Information Criterion [AIC], Comparative Fit Index [CFI], standardized root mean square residual [SRMR]). Competing models were compared to one another using the chi-square difference test as well as the Tucker–Lewis incremental fit index. The latter evaluates improvement in fit for a model of interest as contrasted with a null model. The *LISREL* 8.0 program (Version 8.80, Jöreskog & Sörbom, 2006) was used to conduct the CFAs.

Primary Competing Models Estimated With OSS Data

CFAs were carried out in a stepwise manner in which nested models were systematically evaluated for their fit to the data and their relative fit with respect to each other, taken in succession. Formulation of these models was influenced in part by the EFA results as well as consideration of the psychological and behavioral features under study. Models involving two or three factors allowed the latent factors to be correlated. The models estimated are detailed as follows:

1. A *null model* (where all model parameters were fixed) was estimated that assumed no common latent structure. Although not truly plausible, the null model provides a good baseline against models that do make explicit assumptions regarding latent structure. (A null model is estimated to determine whether or not it can be rejected. There would be no point in modeling a data set in which all variables were uncorrelated.)
2. A *one-factor model* that assumed all features loaded on a single common underlying factor. Such a model is reasonable, as the assessment staff were taking a whole person, holistic approach, which might have yielded a highly interrelated set of variables in the final assessments.
3. A *two-factor model* was formulated that partitioned interpersonal or social and emotional variables (social relations, emotional stability, motivation, energy and initiative, leadership, physical ability, security) from the intelligence processing (effective IQ, propaganda skills, observing and reporting) variables. Such a model allows interpersonal or social features to be linked with emotional stability, holding aside processes clearly linked to cognitive and information processing capabilities. Security is linked to the emotional factors, as it is assumed that one's capacity to maintain security might be reflective of emotional and interpersonal stability. In short, this model is a parsimonious statement that plausibly places personality-related variables together on one factor, whereas those variables more clearly indicative of efficient cognitive and information processing abilities (particularly in relation to the work of an intelligence agent) are on the other factor.
4. A *three-factor model* was formulated, informed both by the EFA results and personality theory, and partitioned the variables into emotional and interpersonal factors (social relations, emotional stability, security), intelligence processing (effective IQ, propaganda skills, observing reporting), and agency/surgency (motivation, energy and initiative, leadership, physical ability). Agency/surgency refers to a personality trait system characterized by energy, motivation, incentive reward, and interpersonal engagement; it has both psychological and physical behavior referents. Separation of the emotional or interpersonal variables from agency/surgency is consistent with modern understanding of personality systems concerning emotion (especially anxiety and fear), affiliation (interpersonal behavior), and positive incentive reward and motivation systems (Depue & Lenzenweger, 2005). Also, in this three-factor model, "motivation for assignment" was positioned with the agency/surgency variables as it represents an incentive reward-oriented, task-oriented posture. For example, many candidates described their primary motivations for an OSS assignment as desire for particular work, active assignment, or leadership as well as a "get the job done" (or, "get the war over with") sentiment.

RESULTS

Exploratory Factor Analysis Results

Table 2 contains the results of the EFA of the OSS data using principal axis factoring with an orthogonal (Varimax)

TABLE 2.—Results of the exploratory factor analysis (EFA) using principal axis factoring (orthogonal rotation).

	Factors		
OSS Variable	Factor 1	Factor 2	Factor 3
Effective IQ	.876		
Propaganda skills	.808		
Observing and reporting	.624		
Social relations		.728	
Emotional stability		.701	.378
Security		.455	
Motivation for assignment		.413	
Energy & initiative			.860
Leadership	.495		.588
Physical ability		.355	.394

Note. Loadings below .35 are suppressed. Extraction method principal axis factoring with orthogonal (Varimax) rotation with Kaiser normalization. OSS = Office of Strategic Services.

rotation. A solution with three factors was retained based on both the Kaiser criterion (eigenvalue ≥ 1.0) and the Scree test (Gorsuch, 1983). Highly similar results to those reported in Table 2 were obtained for the EFA whether using principal axis factoring or, alternatively, principal components analysis. Moreover, the results (number of factors, patterning of loadings) were largely the same whether using an orthogonal (Varimax) or oblique (Oblimin) rotation. What can be seen from Table 2 is that the solution retained is somewhat similar (but clearly not identical) in pattern to the original centroid solution retained by the OSS Assessment staff (see Table 3), but some important differences appear between the two solutions. For example, in the new EFA, Factor 1 consists of effective IQ, propaganda skills, and observing and reporting, but also contains a substantial loading on leadership (which was not present in the original analysis). Factor 2 (Table 2) consists of emotional and interpersonal adjustment items (emotional stability, social relations, security, motivation for assignment), whereas in the original EFA (Table 3) the "adjustment" factor did not load the variable motivation for assignment heavily. Finally, Factor 3 (Table 2) appears to be a factor accounting for agentic or surgent behaviors, consisting of energy and initiative, physical ability, and leadership. It is interesting that the motivation for assignment variable had something of a weak relationship with all of the factors obtained in the original OSS EFA (Table 3), but loads Factor 2 substantially in this analysis.

Interpretation of the original EFA results (Table 3) is hindered by the fact that, as noted earlier, the details of the original factor extraction method (centroid method) and factor rotation method (if any) were not specified in the original report (OSS Assessment Staff, 1948). This information would be considered critical to understanding the results of a factor analysis as reported by today's standards. This missing information is important given some of the striking differences between the modern EFA and the 1948 EFA. For example, as can be seen in the original solution (Table 3), four factors were retained from the analysis of the correlation matrix reported in 1948 and were interpreted by the OSS staff, but only three were retained in the current EFA. It is important to note that even when the complete matrix is analyzed (i.e., including the "overall" variable) using modern software, the results do not support retention of a fourth factor (the fourth factor in the current analysis has an eigenvalue of .77, or well short of the 1.00 that would be needed customarily to justify retention). It is also particularly striking that the factor loadings computed with modern statistical software differ considerably from those reported in the original 1948 analysis (compare loadings in Table 2 with Table 3). That said, the general similarity of the patterning of loadings across the original and current EFA solutions is notable, suggesting some consistency in results. On balance, however, given the lack of detail available for the original analysis, the latent structure of this OSS data set is better understood based on results obtained using modern EFA methods.

Confirmatory Factor Analysis Results

Bearing in mind that the foregoing EFA results reflect a purely empirical, atheoretical exploration of the variables of interest, the theory-guided CFA analyses are considered next. Table 4 contains the actual LISREL-based maximum likelihood factor loadings for the four models described earlier. To evaluate the congruence between a latent structure model and observed data (i.e., goodness of fit), the chi-square statistic is often consulted first and the principle for interpreting it is the

TABLE 3.—Results of the original exploratory factor analysis (EFA) reported in *Assessment of Men* (OSS Assessment Staff, 1948).

	Original Factors Reported			
OSS Variable	Factor 1 Adjustment	Factor 2 Effective Intelligence	Factor 3 Physical Energy	Factor 4 Authoritative Assertion
Effective IQ	−.18	.55	.14	.24
Propaganda skills	.09	.42	−.13	.21
Observing and reporting	.02	.52	−.08	−.07
Social relations	.40	.14	.02	−.15
Emotional stability	.46	−.09	.11	.04
Security	.42	−.02	−.15	−.06
Motivation	.26	−.03	.16	.20
Energy & initiative	.00	.00	.53	.42
Leadership	.03	.14	.35	.39
Physical ability	.05	−.10	.42	−.04
Overall	.14	.43	.06	.05

Note. Factor extraction method = centroid method; rotation (if any) unspecified; diagonal values unspecified. Factor names used were those provided by OSS Assessment Staff (1948). Loadings are reported to two decimal places as in original. "Overall" rating was used in the original analysis and is reported in these results as shown in Table 54 (p. 513) of *Assessment of Men* (OSS Assessment Staff, 1948). The order of the variables in this table is that used in the other tables to facilitate inspection and comparison.

TABLE 4.—Factor loadings for competing models obtained using confirmatory factor analysis.

	Competing Models					
	One Factor	Two Factor		Three Factor		
	1	1	2	1	2	3
OSS Variable	Unifactorial	Emotional/Interpersonal	Intelligence Processing	Intelligence Processing	Emotional/Interpersonal	Agency/Surgency
Effective IQ	.73	—	.94	.99	—	—
Propaganda skills	.62	—	.75	.71	—	—
Observing and reporting	.51	—	.67	.64	—	—
Social relations	.56	.58	—	—	.71	—
Emotional stability	.59	.65	—	—	.89	—
Security	.29	.30	—	—	.42	—
Motivation for assignment	.56	.58	—	—	—	.55
Energy & initiative	.79	.84	—	—	—	.84
Leadership	.84	.83	—	—	—	.85
Physical ability	.31	.39	—	—	—	.35

Note. $n = 133$. — = a LISREL constrained zero loading. These solutions are direct and unique with no rotation necessary. The LISREL program allows one to estimate the degree to which the latent variables underlying the Office for Strategic Services assessment dimensions are correlated in the models containing more than one latent variable (i.e., factor). For the two-factor model the correlation between emotional/interpersonal and intelligence processing = .66 ($p < .001$). For the three-factor model, the correlations were as follows: Intelligence Processing × Emotional/Interpersonal = .29 ($p < .01$); Intelligence Processing × Agency/Surgency = .69 ($p < .001$); and Emotional/Interpersonal × Agency/Surgency = .70 ($p < .001$). These correlations do not reflect a rotation as LISREL solutions are direct as noted.

larger the chi-square value (smaller *p* values), the poorer the fit between the model and data; the smaller the chi-square (large *p* values), the better the fit. Inspection of the chi-square values for the four models (Table 5; null through three-factor) shows a steady decline in magnitude of the chi-square value, suggesting increasingly better fit between the model and OSS data as one moves toward the three-factor model. Similarly, when interpreting both the AIC and SRMR as indexes of fit, the principle is the smaller the AIC and SRMR values, the better the fit. Inspection of the AIC and SRMR values for the four models reveals the three-factor model with the smallest AIC and SRMR values. The SRMR value for the three-factor model (.08) is suggestive of a good fit between a model and the observed data (Hu & Bentler, 1999). The CFI, which contrasts the fit of the model of interest with that provided by the null model, shows that three-factor model provides a generally good fit to the data (CFI = .93), where larger values of the CFI indicate better fit (CFI = 1.00 would indicate a perfect fit). Finally, as noted also in Table 4, LISREL allows one to estimate the degree to which the latent variables (i.e., factors) in the two-factor and three-factor models are correlated. In short, the latent variables in both the two-factor and three-factor model are substantially correlated, however the Intelligence Processing × Emotional/Interpersonal factors are somewhat less strongly associated (although still significantly) as compared to the other factor combinations.

The next step in evaluating the CFA results is to conduct a sequential comparison of models using the differences in the goodness-of-fit chi-square values for the four models. Thus, three contrasts were conducted: (a) the null versus one-factor model, (b) the one-factor versus two-factor model, and (c) the two-factor versus three-factor model. A comparison of the chi-square fit statistics for each model examines the differences in the chi-square values and the degrees of freedom. The difference between these chi-square values is then evaluated for statistical significance. These differences reveal the extent to which one model fits the data better (or worse) than a competing model. To assess the amount of information gained in the comparison of two competing models and to generate an estimate of the improvement in fit obtained in using a better model versus the null model, the non-normed incremental fit index (Tucker–Lewis Index [TLI]; Tucker & Lewis, 1973)

TABLE 5.—Comparison of models estimated with confirmatory factor analysis.

Model	Chi-Square	*df*	*p*	AIC	CFI	SRMR
Model fit						
Null (M_0)	888.07	45	.001	908.07	—	—
One factor (M_1)	249.74	35	.001	289.74	.78	.12
Two factor (M_2)	150.45	34	.001	192.45	.88	.09
Three factor (M_3)	86.30	32	.001	132.30	.93	.08
Model Comparison Statistics						
Contrast	Chi-Square Difference	*df*	*p*	Tucker–Lewis Cumulative		
$M_0 - M_1$	638.33	10	.001	.72		
$M_1 - M_2$	99.29	1	.001	.84		
$M_2 - M_3$	64.15	2	.001	.91		

Note. AIC = Akaike Information Criterion (a stand-alone fit index); CFI = Comparative Fit Index (a stand-alone goodness-of-fit index); SRMR = standardized root mean square residual. Tucker–Lewis cumulative (non-normed fit index) refers to the incremental fit of the proposed model of interest versus the null (independence) model.

was calculated. The results of the model comparisons and the cumulative incremental fit index values are in the bottom panel of Table 5. As can be seen from Table 5, the one-factor model clearly provides a significantly better fit to the data than the null model. However, the two-factor model is a significant improvement in fit over the one-factor model, and the three-factor model is a significant improvement over the two-factor model. The cumulative TLI values reveal that the three-factor model (TLI = .91) is within the realm of a good or acceptable, although not perfect, fitting model given the observed OSS data. Clearly, the stand-alone fit indexes (top panel, Table 5) and the model comparison results (bottom panel, Table 5) point to the three-factor model as providing the best fit to the observed data, with clear superiority over one-factor and two-factor models. It is worth restating that a plausible four-factor model did not fit these data well, but, in fact, generated an invalid solution. Thus, simply adding additional factors does not necessarily improve model fit. In summary, the CFA results strongly suggest that the hypothesized three-factor model described earlier provides the best fit to the 10-variable matrix generated by the original OSS ratings. These results are supportive of the current EFA results and place the proposed three-factor model on a foundation consisting of much firmer statistical information. In short, the OSS assessment team ratings of the candidates reveal three factors at play: intelligence processing, emotional or interpersonal features, and agency/surgency.

Supplementary models estimated with OSS data. In addition to these four primary CFA models, three alternate (but theoretically grounded and plausible) models were estimated in the spirit of analytic thoroughness. One was a two-factor model, the second was a three-factor model, and the final one was four-factor in nature. The alternative two-factor model was one that placed the "security" variable with the intelligence-related (i.e., tradecraft) items. This model was considered as one might think that the ability to keep a secret might depend more closely on skills that covary with the ability to do intelligence-related activities and demonstrate good tradecraft skills (see later). This alternative two-factor model did not fit the observed data as well as the primary two-factor model that placed "security" with the emotional or interpersonal variables ($\chi^2 = 158.44$, AIC = 200.44, CFI = .87, SRMR = .10). Similarly, the alternative three-factor model that also kept "security" with the intelligence-related items did not fit the observed data as well as that found for the primary three-factor model ($\chi^2 = 108.45$, AIC = 154.45, CFI = .91, SRMR = .11). Finally, a four-factor model with the following structure was estimated: Factor 1 (effective IQ, observational skills, propaganda skills) versus Factor 2 (emotional stability, social relations, security) versus Factor 3 (motivation, leadership) versus Factor 4 (energy, physical ability). This model, however, fit the data poorly and generated a Heywood case (Rindskopf, 1984) in the standardized solution results, suggesting the factor solution was invalid.

Discussion

The OSS assessment program was a landmark development for both the intelligence community and the clinical psychological science of personality assessment. In light of what we know today in terms of personality assessment, assessment centers, psychometrics, and statistical analysis, it would be easy enough to criticize the original effort as falling short in one methodological area or another. In fact, it would actually be too easy to do so and would actually risk missing potential value in what was, in fact, accomplished and demonstrated. The importance of the original OSS assessment program for personality assessment at large was described in this journal by Handler (2001) and cannot be overstated. That the OSS program received extensive favorable attention in Wiggins's (1973) classic monograph *Personality and Prediction: Principles of Personality Assessment* for its salience and meaning attests to its foundational status in personality assessment. Although conceived and implemented rapidly nearly 70 years ago amidst profoundly important wartime pressures, it retains a highly unique position in the body of psychological assessment knowledge given its intensity, creativity, and productivity. That individuals such as Murray, Fiske, Lewin, Tolman, Tryon, Gardner, MacKinnon, Miller, Mowrer, Sanford, Symonds, Kluckhohn, Harding, and Bronfenbrenner threw their intellectual powers into the OSS program represents a rare nexus of psychological and assessment talent. All said, the richness of the original set of OSS assessment data could continue to provide useful information and direction to psychological assessment and selection discussions for the intelligence community and related fields, assessment center advocates, and those interested in personality in relation to real-world challenges.

This study tapped into that richness using modern statistical methods to extract more meaning and direction from the original data. The results of this set of factor analyses, both exploratory and confirmatory in nature, as applied to the original OSS assessment data for this sample of candidates, suggest strongly that a three-factor model provides the best overall fit to the data. This result differs from the original report of the OSS Assessment Staff (1948) that presented a four-factor EFA solution to the same data. The difference in the EFA results presented here versus the 1948 EFA results is most likely best explained by the use of a more precise factor solution method and modern computational procedures. In short, this study used principal axis factoring, which is both more precise and reliable than what could be achieved by the centroid method used in the late 1940s. As noted earlier, the centroid method of factor analysis was really more of a rough approximation of what a factor solution might look like for the OSS data, and it was likely used at the time because it was more economical in terms of effort. Moreover, due to the nature of the computational procedures in the centroid approach, the likelihood of human error is elevated (although one must assume care was taken in the original calculations by the OSS Assessment Staff). Nonetheless, modern EFA methods, whether applied to either the 10- or 11-variable matrix, support the finding that a three-factor solution best explains the data. An important extension of this line of analysis was the use of confirmatory methods (CFA), which were completely unavailable in 1948. This powerful, modern statistical methodology allowed for systematic dissection of the OSS data through estimation and comparison of competing substantive or theory-guided models. The CFA results provide strong support for a three-factor model providing the best fit to the observed OSS data. This is important, as EFA methods

can employ arbitrary (Kaiser criterion) or somewhat subjective (Scree test) methods for model selection and rotation to clarify solutions, whereas the CFA approach uses a rigorous statistically principled approach to model selection, and solutions are direct (i.e., do not require rotation). The three-factor model was as follows: Factor I: emotional or interpersonal factors (social relations, emotional stability, security), Factor II: intelligence processing (effective IQ, propaganda skills, observing, reporting), and Factor III: agency/surgency (motivation for assignment, energy and initiative, leadership, physical ability). Additionally, estimation of the associations among the three latent variables (i.e., factors) in this three-factor model suggests positive correlations among the latent variables.

Interpretation of these results requires one to bear in mind the scientific context and climate in psychological science at the time this work was done (mid-1940s). This means that the necessity of maintaining an awareness of what was and was not available to the assessors in the OSS assessment program by way of theoretical models and empirical research in assessment cannot be overstated. As noted previously, aside from Murray's (1938) model of personality, the field of clinical psychology (itself in its infancy) had very little to draw on by way of theory. Classical Freudian psychoanalysis and the principles of Skinnerian operant conditioning provided little leverage with respect to systematic nomothetic personality and psychopathology assessment. Moreover, there were no factorial models of normal personality for use as a basis for assessment. Thus, whereas today one is tempted to think of any of the prominent factorial models (e.g., the popular "five-factor"/Big Five model) of normal personality as forming a basis for assessment and selection, such models were simply not available at the time. Moreover, as noted by the OSS Assessment Staff (1948), the nature of the selection task, the novelty of the assessment center approach, the time pressures and urgency placed on the assessment staff owing to the war, and the desire to be both comprehensive and creative in assessment compelled the development of a highly original approach to assessment of OSS candidates. The OSS assessors were not seeking to test a theoretical model or be constrained by a particular substantive approach; rather, they sought to make the best assessment decisions they could make under the circumstances. One could easily argue that the absence of a constraining a priori model driving the assessment approach was a strength of the overall effort. Finally, one must bear in mind the nature of the men and women who were the subjects of the OSS assessment program. In short, they were self-selected in some sense (i.e., willing to be considered for OSS service) and that likely reduced the diversity in the overall pool somewhat, as well as potentially establishing a relatively high floor on some psychological, behavioral, and other characteristics.

In this context it is worth noting that the final scores for each candidate on each of the 10 dimensions were done by the OSS assessment staff as ratings. Although they were "clinical ratings," they were ratings based on an exceptionally rich amalgam distilled from multiple diverse data streams. They were derivative from results coming from psychological tests, clinical interviews, group interactions, performance on behavioral tasks, case history information, physical agility tasks, and so forth. Thus, they were not akin to typical symptom ratings based solely on a structured clinical interview instrument as applied by an assessor in a one-on-one interview with a research or clinical subject. Nor were they akin to descriptive statements made about a person based on observed levels on psychological dimensions assessed in a typical psychometric personality inventory. Rather, the ratings made by the OSS assessment staff made use of all available data on the candidates. The approach embodied in spirit an assessment perspective that would eventually be recommended nearly 40 years later, the so-called LEAD (longitudinal, expert clinicians, all available data) standard approach advocated by Spitzer (1983), for use in psychological and psychiatric assessments. Finally, in a related vein, although the "rating" (as opposed to "counting") approach to psychological assessment is imperfect, there remain today (just as in the 1940s) no definitive ratio scale-based approaches to psychological (or personality) adjustment assessment and there are no biologically based psychological (or personality) parameters that can unequivocally claim superiority to the clinical ratings approach.

Moving on to an evaluation of the obtained CFA results, each factor is considered in turn. Factor I in the three-factor CFA model (see Table 4) reveals a cohesive structure linking effective or practical intelligence (in the IQ sense) to tradecraft skills. The abilities to observe carefully, take note of important information, report that information back as appropriate, or use it in the service of propaganda efforts are clearly connected to what the original OSS Assessment Staff termed effective intelligence (really, IQ that can be put to use). In other words, to be able to think on one's feet and to attend to patterns, developments, or puzzling aspects of a situation clearly appears to have been strongly associated with native intelligence as gleaned from candidates by the OSS assessors. The results reported here strongly support that association between tradecraft potential and effective intelligence as well.

The second factor of this model (Factor II, Table 4) points to important personality systems discernible within these OSS candidate ratings data. The social relations aspect of the factor points to the salience of interpersonal skills, long known to be a critical component of personality, as viewed by the OSS assessors. Although named differently in various personality models or systems, the interpersonal component of personality is linked to the affiliation system (Depue & Lenzenweger, 2005) and is reflective of one's ability to interpersonally connect with and interact with others, often termed *communion* (Wiggins, 1991). (Another personality factor relevant to interpersonal behavior emerged in Factor III and is discussed later.) Emotional stability, also a component of Factor II, represents the interaction of the constraint and control system within human personality in relation to both the positive (e.g., joy, happiness) and negative (e.g., fear, anxiety, anger) emotion systems (see Depue & Lenzenweger, 2005, for review). That security concerns loaded with the emotional or interpersonal (Factor II) aspects of the OSS data speaks to the likelihood that one's ability to bluff and maintain a cover are associated with the ability to control emotions and behavior in the service of security. This connection was also seen and remarked on by the original OSS assessment staff. These modern analyses support linking security concerns with emotional stability and interpersonal relations and the consideration of alternative CFA models (noted earlier) strengthens it. A practical implication of Factor II concerns the signal it sends to attend to emotional and interpersonal stability factors in

relation to security with respect to personnel selection even today.[3] More specifically, for example, one can extend this concern to an assessment focus on forms of interpersonal dysfunction, particularly the personality disorders, which deserve careful scrutiny in candidate assessments. Personality disorders represent that domain of psychopathology that can be harbored within a seemingly normal appearing personality, but that exert their impact through impaired judgment, dysfunctional interpersonal behaviors, and emotional dyscontrol. Finally, in discussing Factor II, it is worth describing the results of a highly informal survey conducted by the author of a subset of surviving OSS officers (Lenzenweger, 2011, unpublished raw data) as to what they thought, in retrospect 65 years after the war, were attributes of a successful OSS officer. Many of those surveyed mentioned two important features—lack of fear and an ability to stay in control and not panic. Both of these features mentioned by former OSS officers clearly fall within this factor of emotional stability and social relations.

Factor III (energy and initiative, leadership, physical ability, motivation for assignment) of this model points clearly to another major personality system, namely the system connected to positive emotion (energy, outgoingness, vitality) that derives from what is known as the incentive motivation or approach/reward system (see Depue & Lenzenweger, 2005, for review). The behavioral characteristics of individuals who have strong positive emotion or incentive reward profiles are marked by a sense of surgency or agency (Wiggins, 1991), sometimes referred to as *agentic extraversion* (assertiveness, dominance, outgoingness, decisiveness). In short, they tend to engage with the world, seek rewards, display a sense of potency, and seek rewarding goals. The neural system underpinning this set of psychological and behavioral characteristics has been described in great detail (Depue & Lenzenweger, 2005). The configuration of features loading on Factor III also shares elements of what is termed *psychological hardiness* (Bartone, Roland, Picano, & Williams, 2008) in that persons elevated on hardiness are strongly committed, actively engaged in their environment, enjoy new situations and challenges, and are internally motivated as well as creating their own sense of purpose (see Bartone et al., 2008, p. 78). The implication of this particular factor (Factor III) for personnel assessment and selection is especially strong; namely, one seeks energetic, outgoing, internally motivated, and vibrant people to take on challenging and potentially rewarding tasks such as those afforded by intelligence activities, whether analytic or operational. At the same time, this particular factor points to a clinical consideration in that one would necessarily want to be attentive to extremes in this profile. Such extremes could yield, in some circumstances, unduly self-centered, even pathological, postures with respect to others and an abnormally high level of self-serving internal motivation. An example of the latter could be found in those people termed clinically narcissistic or as evidencing malignant narcissism (see Kernberg, 1984; Post, 2003). Another dysfunctional manifestation of an extreme level of Factor III features could be psychopathy as evidenced by excessively high levels of boldness, meanness, and disinhibition (Patrick, Fowles, & Krueger, 2009). In such dysfunctional forms, dominating aggressive behavior directed at others is easily manifested and associated with deficits in social empathy, impaired identity formation, and psychological maneuvers that place blame on others (e.g., Lenzenweger, McClough, Clarkin, & Kernberg, 2012). Clearly, assessment and screening efforts seeking well-adapted individuals with Factor III features would need also to incorporate checks to help to ensure against retention of potential candidates who manifest pathologically narcissistic or psychopathic features.

As with any study, there are a number of specific caveats to keep in mind when analyzing these results. First and foremost, it would have been optimal to have had a much larger sample size for these factor analyses. As noted earlier, more than 5,000 candidates were assessed at Station S and other sites; however, the data from those assessments were reported as having been destroyed soon after the end of World War II (Handler, 2001). If even the OSS summary sheets (Figure 1) had been retained, a larger sample size could have been assembled from such raw data for statistical analysis. Nonetheless, the subject to variable ratio for this study was acceptable and the sample size of 133 is in the range associated with empirical studies that can recover factor structures reliably from observed data (Preacher & MacCallum, 2002). Notwithstanding these caveats, the OSS data set used here is extraordinarily unique and affords a glimpse into one of the most ambitious and important assessment projects ever undertaken; thus, analysis of the available data made good sense.

Finally, a word about statistical methods and meaning extracted from results is in order. Factor analysis is only a statistical procedure—or, more accurately, a family of procedures—and the results of such statistical analyses are not self-interpreting. Therefore, despite the technical nature of this endeavor, the results have required interpretation and have been interpreted in light of contemporary personality and psychopathology constructs. The interpretation of the results has necessarily been kept brief. However, it is possible that others might have sought to link these specific findings to other similar personality constructs used in other models. This is not the forum for an extended discussion of alternative conceptualizations of all personality constructs. That said, the results discovered, and the psychological constructs used to understand those results, do provide interesting clues to those components or constructs that are worthy of inclusion in any assessment approach used in the selection of personnel. Clearly, it is possible that other personality constructs might be considered as well (see Lenzenweger et al., 2012).

In summary, this study used modern statistical methods to revisit a highly rarefied correlation matrix that summarized an immense amount of information gained through the assessments done at Station S by the OSS Assessment Staff (1948). There are points of convergence between the results of the initial factor analysis of these data done in 1948 and those results reported here. However, the present results have the benefit of

[3] In considering the potential utility of these models in terms of their constituent constructs and how these constructs might be of use today in assessing and selecting individuals for intelligence services (including clandestine operations) and special operations, one should clearly consider the possibility that the relationship of the factors in these models to success in performance might be nonlinear. That is, it might be that there is a best fit on each (or an optimal combination of the three) factor(s) that is related to success in functioning in target roles. This important observation was generously shared by an anonymous reviewer of this article.

greater reliability and precision in terms of the EFA approach. The CFA approach used here, moreover, provided a powerful approach to the OSS data that enabled this study to home in on a model that provides the best fit to these unique data. The three-factor model presented here might be useful in other discussions in the intelligence community where personnel selection issues are always of great importance. The technical aspects of this study provided the tools to discern structure in the OSS data, and the results point to the importance of the following domains for selection: IQ and intelligence-related skills, emotional and interpersonal stability, and agency/surgency.

Acknowledgments

I thank Charles Pinck of *The OSS Society* of Falls Church, VA, for his gracious assistance in facilitating access to former members of the OSS. I thank former members of the OSS for their time and energy in providing valuable insights regarding their experiences in the OSS. I thank John Harding, PhD (late), and Urie Bronfenbrenner, PhD (late), former members of the assessment staff at Station S, and John L. Behling, PhD, former OSS officer, for helpful discussions. I am also grateful to L. Morgan Banks, PhD, for helpful comments on an early version of this article.

References

Banks, L. M. (1995). *The Office of Strategic Services psychological selection program* (Unpublished master's thesis). U.S. Army Command and General Staff College, Fort Leavenworth, KS.

Banks, L. M. (2006). The history of special operations psychological selection. In A. D. Mangelsdorff (Ed.), *Psychology in the service of national security* (pp. 83–95). Washington, DC: American Psychological Association.

Bartone, P. T., Roland, R. R., Picano, J. J., & Williams, T. J. (2008). Psychological hardiness predicts success in US Army Special Forces candidates. *International Journal of Selection and Assessment*, *16*, 78–81.

Butcher, J. N. (2010). Personality assessment from the nineteenth to early twenty-first century: Past achievements and contemporary challenges. *Annual Review of Clinical Psychology*, *6*, 1–20.

Byham, W. C. (2002). *What is an assessment center? The assessment center method, applications, and technologies*. Pittsburgh, PA: Development Dimensions International.

Depue, R. A., & Lenzenweger, M. F. (2005). A neurobehavioral model of personality disturbance. In M. F. Lenzenweger & J. F. Clarkin (Eds.), *Major theories of personality disorder* (2nd ed., pp. 391–453). New York, NY: Guilford.

Gorsuch, R. L. (1983). *Factor analysis* (2nd ed.). Hillsdale, NJ: Erlbaum.

Handler, L. (2001). *Assessment of men:* Personality assessment goes to war by the Office of Strategic Services Assessment Staff. *Journal of Personality Assessment*, *76*, 556–578.

Harman, H. H. (1976). *Modern factor analysis* (3rd ed., rev.). Chicago, IL: University of Chicago Press.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*, 1–55.

Hyams, K. C., Jones, E., & Wessely, S. (2003). Screening for vulnerability to psychological disorders in the military: An historical survey. *Journal of Medical Screening*, *10*, 40–46.

Jöreskog, K. G., & Sörbom, D. (2006). LISREL 8.8 for Windows [Computer software]. Skokie, IL: Scientific Software International.

Kernberg, O. F. (1984). *Severe personality disorders*. New Haven, CT: Yale University Press.

Lenzenweger, M. F. (2011). [Recollections of surviving OSS officers]. Unpublished raw data.

Lenzenweger, M. F., Dworkin, R. H., & Wethington, E. (1989). Models of positive and negative symptoms in schizophrenia: An empirical evaluation of latent structures. *Journal of Abnormal Psychology*, *98*, 62–70.

Lenzenweger, M. F., McClough, J. F., Clarkin, J. F., & Kernberg, O. F. (2012). Exploring the interface of neurobehaviorally-linked personality dimensions and personality organization in borderline personality disorder: The Multidimensional Personality Questionnaire and Inventory of Personality Organization. *Journal of Personality Disorders*, *26*, 902–918.

MacCallum, R. C., Widaman, K. F., Zhang, S., & Hong, S. (1999). Sample size in factor analysis. *Psychological Methods*, *4*, 84–99.

Mackinnon, D. W. (1980). *How assessment centers were started in the United States: The OSS assessment program*. Pittsburgh, PA: Development Dimensions International. (Original work published 1974)

Murray, H. A. (1938). *Explorations in personality*. New York, NY: Oxford University Press.

Office of Strategic Services Assessment Staff. (1948). *Assessment of men: Selection of personnel for the Office of Strategic Services*. New York, NY: Rinehart.

Patrick, C. J., Fowles, D. C., & Krueger, R. F. (2009). Triarchic conceptualization of psychopathy: Developmental origins of disinhibition, boldness, and meanness. *Development & Psychopathology*, *21*, 913–938.

Pols, H., & Oak, S. (2007). War and military mental health: The US psychiatric response in the 20th century. *American Journal of Public Health*, *97*, 2132–2142.

Post, J. M. (2003). Assessing leaders at a distance: The political personality profile. In J. M. Post (Ed.), *The psychological assessment of political leaders* (pp. 69–104). Ann Arbor: University of Michigan.

Preacher, K. J., & MacCallum, R. C. (2002). Exploratory factor analysis in behavior genetics research: Factor recovery with small sample sizes. *Behavior Genetics*, *32*, 153–161.

Rindskopf, D. (1984). Structural equation models: Empirical identification, Heywood cases, and related problems. *Sociological Methods and Research*, *13*, 109–119.

Spitzer, R. L. (1983). Psychiatric diagnosis: Are clinicians still necessary? *Comprehensive Psychiatry*, *24*, 399–411.

Strecker, E. A. (1944). Military psychiatry: World War I, 1917–1918. In J. K. Hall (Ed.), *One hundred years of American psychiatry* (pp. 385–416). New York, NY: Columbia University Press for American Psychiatric Association.

Tucker, L. R., & Lewis, C. A. (1973). A reliability coefficient for maximum likelihood factor analysis. *Psychometrika*, *38*, 1–10.

Waller, D. (2011). *Wild Bill Donovan: The spymaster who created the OSS and modern American espionage*. New York, NY: Free Press.

Wiggins, J. S. (1973). *Personality and prediction: Principles of personality assessment*. Reading, MA: Addison-Wesley.

Wiggins, J. S. (1991). Agency and communion as conceptual coordinates for the understanding and measurement of interpersonal behavior. In D. Cicchetti & W. M. Grove (Eds.), *Thinking clearly about psychology: Vol. 2. Personality and psychopathology* (pp. 89–113). Minneapolis: University of Minnesota Press.

Williams, T. J., Picano, J. J., Roland, R. R., & Bartone, P. (2012). Operational psychology: Foundation, applications, issues. In J. H. Laurence & M. D. Matthews (Eds.), *The Oxford handbook of military psychology* (pp. 37–57). New York, NY: Oxford University Press.

Woodworth, R. (1919). Examination of emotional fitness for war. *Psychological Bulletin*, *15*, 59–60.