

International Workshop on Military Recruitment and Retention in the 21st century

Psychological tests for personnel selection in Estonian Defence Forces:
New tools and new areas of utilization

Aaro Toomela
Academy of the Estonian Defence Forces
Department of Social and Behavioural Sciences
Riia 12, Tartu, 51013, Estonia

E-mail: aaro.toomela@ksk.edu.ee

& Harri Ints
General Staff of the Estonian Defence Forces
Narva maantee 8, Tallinn, 15007, Estonia

E-mail: Harints@solo.delfi.ee

Abstract

Estonian Defence Forces started to develop a new psychological test battery for recruitment and personnel selection in 1999. A preliminary test battery that covered most of the psychological characteristics of healthy persons (cognitive abilities, personality, collectivism, coping styles, conceptual structure, machiavellism, etc.) was created. Some of the measures have never before been applied for such purposes. The validity of the battery was assessed in about 200 recruits and 200 career officers. All participants were studied twice. First participants were tested with the test battery. 6-8 months later, military performance of participants was measured.

Participants were divided into three groups according to the military performance level, inappropriate, appropriate, and successful. It was possible to classify 76% of participants into correct performance groups on the basis of psychological test performance. In the Discriminant Function Analysis, novel psychological measures (conceptual structure, “situational” form of a personality questionnaire) gave a statistically independent and significant contribution to the prediction of performance groups.

The same test battery was also used for predicting other characteristics of participants. Participants were tested twice with the Beck Depression Inventory to describe the level of depression of participants. In most cases the number of symptoms for depression decreased during service. There was a subgroup of recruits, however, whose number of depression symptoms increased considerably. Eighty-six percent of those were correctly classified on the basis of the results of psychological test battery. Thus, psychological tests allow identifying persons with a risk for depression and planning appropriate actions to prevent the increase of depression in such persons. In a similar way we were able to identify at about 70% accuracy which attitudes persons have towards alcohol and narcotics.

Acknowledgements: This work was supported by the Grant No. 01/2000 from the Estonian Ministry of Defence.

Personnel selection in military environments very often relies on performance of recruits or career officers on a battery of psychological tests. The history of using such tests for predicting military performance dates back to WW1 when US Army introduced psychological testing in 1917 (Yerkes, 1921). What was measured then was basically intelligence. Different countries have developed their own psychological testing systems since then. Measurement of intelligence seems to be a core of such tests in many cases. Some countries, in addition, use additional measures for personality or some other characteristics. A form of measurement may also vary to some degree. In addition to usual paper-and-pencil tests interviews are conducted in some countries (Sweden, for example).

Psychological testing for personnel selection has also moved out from military environment to non-military organizations as well. Despite very broad area of usage, it seems to us, there is no universally acceptable *theory* on what characteristics of psyche are important for a successful career. Many different qualities can be measured by psychological tests. And an impressive number of tests have been developed in the history of psychology (there are thousands of tests available for neuropsychological testing alone, cf. Lezak, 1995).

So, there are thousands of psychological tests measuring many different qualities of the mind. A selection, what specifically should be used for the best prediction of career success should be made. That selection should be supported by a sound theory – which may not exist. Nevertheless, test batteries for personnel selection are usually built as if there is such a theory – only a very limited number of characteristics are usually measured (intelligence, personality).

Estonia, after the collapse of the Soviet Empire, started to build its own defence forces. It was decided in 1998 that personnel selection in the defence forces should be partly guided by psychological test results. So, we started to develop a new test battery for military personnel selection in 1999. As we were not able to find a good theory for helping to decide what are the characteristics that should be measured by a test batter we composed a very large battery which contains subtests for measuring very different characteristics. Here are some examples (special qualities measured are in italics):

NEO-PI (Five-Factor-Model)	<i>Attitudes towards alcohol and narcotics</i>
<i>Neuroticism</i>	<i>Level of Depression</i>
<i>Extraversion</i>	Coping styles
<i>Openness</i>	<i>Problem/task oriented</i>
<i>Agreeableness</i>	<i>Social/emotional</i>
<i>Conscientiousness</i>	<i>Avoidance</i>
FFM – “Situational” (Toomela-Pulver-Valsiner)	Aggressiveness
<i>Neuroticism</i>	<i>Verbal</i>
<i>Extraversion</i>	<i>Physical</i>
<i>Openness</i>	<i>Anger</i>
<i>Agreeableness</i>	Values
<i>Conscientiousness</i>	<i>Individual</i>
+ <i>Attitudes towards alcohol and narcotics</i>	<i>Community</i>
<i>Mental abilities (IQ)</i>	Relationship styles
<i>Conceptual Structure</i>	<i>Warmth</i>
Collectivism (3-Factor-Model)	<i>Status</i>
<i>Family</i>	<i>Dependence</i>
<i>Peers</i>	<i>Extravert</i>
<i>Society</i>	<i>Sensitivity for feelings</i>

Some of the tests were new, specially designed for the inclusion in the battery. Recent developments in neuropsychology, developmental psychology and cultural psychology led us to understanding that there are some mental qualities that rarely are studied. A school of

cultural-historical psychology, founded by Lev Vygotsky, proposed that one fundamental human characteristic is the way *how* information is processed. That way of information processing is reflected in the conceptual structure. Different types of conceptual structure can be found. Most frequent in modern cultures are so-called “complexes” and “scientific” concepts. In complexes the basis for information processing is knowledge acquired from observation of everyday activities, in “scientific” concepts the structure of a concept is hierarchical, one attribute or a group of attributes, defined on the basis of experimentation and formal-logical thinking, constitutes a core of a concept. The evidence for individual differences in conceptual structure comes from different fields of studies (cf. Kikas, in press; Luria, 1974; Luria, 1979; Toomela, in press; Toomela, Tomberg, Orasson, Tikk, & Nõmm, 1999; Vygotsky & Luria, 1994; Vygotsky, 1996; Vygotsky & Luria, 1930).

Theoretically, individual differences in conceptual structure should be related to individual differences in every semiotically mediated mental operation. It should be mentioned, that, theoretically, majority of human mental operations are semiotically mediated (Toomela, 1996; Toomela, 2000; Toomela, in press). Thus, we constructed a test for measuring conceptual structure. The test is adapted from Luria (1979). Briefly, that test includes three ways for describing a conceptual structure, definition, categorization, and finding of similarities. In “scientific” concepts links between words are made between items at different hierarchical levels (e.g., 'pets' -- dog, parrot); in complexes links between items at the same level of taxonomic hierarchy are represented (e.g., dog-cow).

In addition, in cooperation with Aleksander Pulver (Academy of the Estonian Defence Forces) and Jaan Valsiner (Clark University, MA), we developed a new version of a personality test. Based on the same theoretical ideas, we hypothesized that usual way of building items in personality tests requires “scientific” thinking to answer questions meaningfully. These items are defined in very general terms with no direct reference to everyday situations. Thus, the answers to those items by persons whose leading conceptual structure is “complex”, we may get very misleading information. We developed a test where 10 items from NEO-Personality Inventory (Costa & McCrae, 1985, Estonian version by Pulver, Allik, Pulkkinen, & Hämäläinen, 1995), 2 from each of the five factors, were used. We put the same item into three different situations. So, we got a test where there were 6 items for every personality factor.

We expected that broad range of tests allows us to develop a test battery with good predictive power for military career. In addition, even though it might be possible that the predictive power of the test battery does not exceed that achieved by other countries, we can get also closer to understanding what specific qualities make a person appropriate for a successful career. We hypothesised that our novel tests as well as tests that usually are not included in the personnel selection test batteries may give significant and independent contribution to the prediction of the career success.

Empirical Study of the Predictive Power of a Test Battery

We conducted a study where recruits and career officers were studied twice – first with the test battery and second time, about 6-8 months later with an inventory describing success in military service (knowledge, opinions of peers and supervisors, etc.). Some of the results of this study are reported next.

First, we estimated the predictive power of our test battery regarding success in a military career. Results for recruits are only reported here. We divided participants into three groups according to their performance in service. Results of different psychological tests were used as predictors. Discriminant Function Analysis revealed that it was possible to classify persons statistically significantly above chance (Wilks' Lambda: .57 approx.

$F(62,374)=1.94$ $p < .0001$). Almost 76% of participants were correctly classified. Details of classification and variables that contributed statistically significantly for the prediction are given in Table 1 and Table 2, respectively. How exactly statistically significant predictors distinguished between groups, is described in Table 3.

Table 1. Classification of participants by success in military service

Observed	Percent Correct	Predicted		
		Inappropriate	Appropriate	Successful
Inappropriate	46.67	14	16	0
Appropriate	91.34	4	137	9
Successful	40.00	0	24	16
Total	75.90	18	177	25

Table 2. Statistically significant variables in the model

	Wilks' Lambda	F-remove (2,187)	p-level
Neuroticism (Situational)	.60	4.65	.011
Conceptual structure	.59	3.98	.020
Openness (NEO-PI)	.59	3.04	.049
Attitudes tow. alcohol (Situational)	.59	2.99	.052
Collectivism – peers	.58	2.73	.068

Table 3. A type of relationship between significant predictors and career success

Neuroticism – decreases	→ successful
Conceptual structure – hierarchical/“scientific”	
Openness – increases	
Attitudes – less tolerant	
Collectivism towards peers – decreases	

It can be seen from the results that, indeed, novel measures contributed significantly for the prediction of military career success. It is also noteworthy, that predictive power was quite high. If the test results were used for personnel selection none of the successful recruits would have been classified as inappropriate, and none of the inappropriate recruits would have been classified as highly successful.

Next, we estimated the predictive power of our test battery regarding the change in the level of depression during military service. Results for recruits are only reported here. The number of symptoms was estimated twice in our study. We divided participants into three groups according to difference score between first and second measurement. Overall, the average level of depression decreased during service. One group comprised persons whose level of depression decreased 1SD more than average, second group comprised persons whose level of depression had an average (+/- 1 SD), and the third group comprised persons whose level of depression increased significantly during service. Results of different psychological tests were used as predictors. Discriminant Function Analysis revealed that it was possible to classify persons statistically significantly above chance (Wilks' Lambda: .33 approx. $F(62,510)=6.02$ $p < 0.00001$). Almost 90% of participants were correctly classified. Details of classification and variables that contributed statistically significantly for the

prediction are given in Table 4 and Table 5, respectively. How exactly statistically significant predictors distinguished between groups, is described in Table 6.

Table 4. Classification by change in the level of depression

Observed	Percent Correct	Predicted		
		High decrease	Average decrease	High increase
High decrease	45.45	10	11	1
Average decrease	95.74	3	225	7
High increase	74.19	0	8	23
Total	89.58	13	244	31

Table 5. Statistically significant variables in the model

	Wilks' Lambda	F-remove (2,255)	p-level
Level of depression (Time 1)	.58	95.06	0.000
Coping style – Avoidance	.35	6.02	.003
Conscientiousness (NEO-PI)	.34	3.69	.026
Aggressiveness – Verbal	.34	3.47	.033
Conceptual structure	.34	3.41	.034
Collectivism – Society	.34	3.35	.036
Openness (Situational)	.34	2.38	.094

Table 6. A type of relationship between significant predictors and change in depression level

Level of depression (Time 1) – high	➔ Increase in depression
Openness (Situational) – lowest	
Conscientiousness (NEO-PI) – highest	➔ Average decrease
Aggressiveness – Verbal – highest	
Conceptual structure – highest “scientific”	
Collectivism – Society – highest	
Coping style – Avoidance – lowest	

It can be seen, again, that test battery has substantial predictive power and novel measures contribute significantly to the prediction. It is also noteworthy, that variables that gave most significant contribution to the prediction of military career success do not overlap with those that predict change in depression level. Prediction of change in depression can be very useful for identifying persons who need some psychological support during service for avoiding the early dropout and increasing the level of retention. Thus, it might be useful to compose psychological test batteries not only for predicting career success but also for predicting other characteristics important for retention.

Finally, we estimated the predictive power of our test battery regarding attitudes towards narcotics and alcohol. Results for recruits are only reported here. We divided participants into three groups according to the attitudes measured by a special inventory that was included in our test battery. Results of different psychological tests were used as

predictors. Discriminant Function Analysis revealed that it was possible to classify persons statistically significantly above chance (Wilks' Lambda: .66 approx. $F(58,530)=2.09$ $p < .0001$). Almost 70% of participants were correctly classified. Details of classification and variables that contributed statistically significantly for the prediction are given in Table 7 and Table 8, respectively. How exactly statistically significant predictors distinguished between groups, is described in Table 9.

Table 7. Classification by attitudes towards alcohol and narcotics

Observed	Percent Correct	Predicted		
		Nontolerant	Average	Tolerant
Nontolerant	30.3	20	45	1
Average	89.6	13	173	7
Tolerant	35.1	1	23	13
Total	69.6	34	241	21

Table 8. Statistically significant variables in the model

	Wilks' Lambda	F-remove (2,265)	p-level
Collectivism – Family	.69	6.54	.002
Relationship style – Status	.68	5.09	.007
Values – individual	.68	3.35	.036
Conceptual structure	.67	2.65	.072
Neuroticism (NEO-PI)	.67	2.33	.099
Conscientiousness (Situational)	.67	2.32	.099

Table 9. A type of relationship between significant predictors and attitudes towards narcotics and alcohol

Collectivism – Family – lowest	→ tolerant
Relationship style – Status – lowest	
Values – individual – lowest	
Conceptual structure – lowest	
Neuroticism (NEO-PI) – highest	
Conscientiousness (Situational) – lowest	→ average

Results of this analysis demonstrate that the same test battery can be used for very different purposes. Prediction of attitudes towards narcotics and alcohol can be helpful for designing programs for fighting against the development of addiction during military service as well as after that.

Conclusions

- (1) The same battery of tests can be useful for predicting very different characteristics of recruits, including success in a military career, increase in the level of depression during service, tolerant attitudes towards narcotics and alcohol.
- (2) The way of information processing – leading type of a conceptual structure – contributes significantly to prediction of different characteristics.
- (3) Testing for military purposes would rely more on developments in cultural psychology, neuropsychology, and developmental psychology.

Reference List

- Costa, P. T., & McCrae, R. R. (1985). The NEO personality inventory manual. Odessa, FL: Psychological Assessment Resources.
- Kikas, E. (in press). Constructing knowledge beyond senses: Worlds too big and too small to see. In A. Toomela (Ed.), Cultural guidance in the development of the human mind. Greenwood.
- Lezak, M. D. (1995). Neuropsychological assessment. Oxford: Oxford University Press.
- Luria, A. R. (1974). Ob istoricheskom razvitii poznavatel'nykh processov. Eksperimental'no-psikhologicheskije issledovanije. Moscow: Nauka.
- Luria, A. R. (1979). Jazyk i soznanije. Moscow: Izdatel'stvo Moskokogo Universiteta.
- Pulver, A., Allik, J., Pulkkinen, L., & Hämäläinen, M. (1995). A Big Five personality inventory in two non-Indo-European languages. European Journal of Personality, 9, 109-124.
- Toomela, A. (1996). How culture transforms mind: A process of internalization. Culture and Psychology, 2(3), 285-305.
- Toomela, A. (2000). Activity theory is a dead end for cultural-historical psychology. Culture and Psychology, 6(3), 353-364.
- Toomela, A. (in press). Culture as a semiosphere: On the role of culture in culture-individual relationship. In I. E. Josephs, & J. Valsiner (Eds.), Dialogicality in development (p. xxx-xxx). Elsevier.
- Toomela, A., Tomberg, T., Orasson, A., Tikk, A., & Nõmm, M. (1999). Paradoxical facilitation of a free recall of nonwords in persons with traumatic brain injury. Brain and Cognition, 39, 187-201.
- Vygotsky, L., & Luria, A. (1994). Tool and symbol in child development. (Originally written in 1930). In R. van der Veer, & J. Valsiner (Eds.), The Vygotsky reader. (pp. 99-174). Oxford, UK: Blackwell.
- Vygotsky, L. S. (1996). Myshlenije i rech. (Thinking and speech. Originally published in 1934). Moscow: Labirint.
- Vygotsky, L. S., & Luria, A. R. (1930). Etjudy po istorii povedenija. Obezjana. Primitiv. Rebjonok. Moscow-Leningrad: Gosudarstvennoje Izdatel'stvo.
- Yerkes, R. M. (Ed.). (1921). Psychological examining in the Unites States army. Memoirs of the National Academy of Sciences, Vol. XV. Washington: Government Printing Office.