SAFETY ASSESSMENT RAIL

Short name SARAIL

Version 01

Mödling, July 2012
Copyright © 2012 by SCHUHFRIED GmbH
Author of the test set SCHUHFRIED GmbH
Author of the manual SCHUHFRIED GmbH
CONTENTS

1 INTRODUCTION AND OVERVIEW ....................................................... 3

2 APPLICATION .................................................................................. 4
  2.1 Dimension selection ................................................................... 4
  2.2 Construction of the test set ......................................................... 5
  2.3 Psychometric properties ............................................................. 7

3 DIMENSIONS ................................................................................. 10
  3.1 Cognitive abilities ..................................................................... 10
  3.2 Stress tolerance and ability to react ............................................ 11
  3.3 Attention and concentration ...................................................... 12
  3.4 Sensomotor functions and visual perception ............................... 12
  3.5 Determinants of personality ....................................................... 13

4 SCORING AND INTERPRETATION .................................................. 15
  4.1 Results on each dimension ......................................................... 15
  4.2 Standardization ......................................................................... 15
  4.3 Report ......................................................................................... 16
  4.4 Case studies .............................................................................. 16

5 REFERENCES .................................................................................. 19

6 INDEX OF TABLES AND ILLUSTRATIONS ................................. 21
1 INTRODUCTION AND OVERVIEW

In the Vienna Test System the SCHUHFRIED company has for over 60 years been providing high quality and scientifically based tests for use in human resources. The test set Safety Assessment Rail (SARAIL) forms part of the Vienna Test System Human Resources (Vienna Test System HR). A test set is a combination of test dimensions that addresses a particular complex assessment issue. A test set relieves you of the task of selecting tests and enables you to be certain of using the right dimensions for the particular assessment situation. In addition, the test set simplifies cross-test evaluation and interpretation for you and hence makes it easier to answer the question being addressed.

The SARAIL test set was developed within the Vienna Test System HR to provide a specific tool for economic and high-quality selection and ongoing monitoring of train drivers. As a first step towards finding your way around, Figure 1 summarizes the information contained in this and other manuals:

![Diagram](image)

Figure 1: Overview – manuals
2 APPLICATION

The SARAIL test set is a psychological assessment package that has been developed within the framework of the Vienna Test System HR for the assessment of train drivers’ fitness to drive. Psychological tests are used in the selection and continuous monitoring of the psychological aptitude of train drivers in order to ensure that applicants and existing staff have the capacity to drive a train safely. The test set comprises tests to assess (1) cognitive abilities, (2) stress tolerance and ability to react, (3) attention and concentration, (4) sensomotor functions and perception and (5) relevant personality dimensions.

The EU’s Directive 2007/59/EC of the European parliament and of the council of 23 October 2007 on the certification of train drivers operating locomotives and trains on the railway system in the Community stipulates that would-be train drivers must undergo psychological assessment and that active train drivers must also be tested periodically. SARAIL is used to test not only train drivers but also drivers of other railed vehicles (e.g. rapid-transit system, metro/underground, tram) and people in safety-related jobs in the rail sector.

The aim of the SARAIL test set is to improve the selection and monitoring of train drivers in terms of both economy and quality by making use of state-of-the-art technological developments. The SARAIL test set provides an economical means of obtaining important information about driver characteristics that cannot be reliably obtained using other suitability assessment instruments. It therefore makes an enormous contribution to the reliability of the decisions that are made.

The selection of dimensions for the SARAIL test set is based on EU Directive 2007/59/EC and the criteria of the CER (Community of European Railways) and the ETF (European Transport Workers’ Federation) (see Section 2.1).

2.1 Dimension selection

EU Directive 2007/59/EC stipulates that both trainees and existing train drivers must undergo psychological assessment. As a minimum, testing of train drivers must cover the following dimensions:

- logical reasoning
- memory
- ability to react
- attention and concentration
- perception
- eye-hand coordination
- relevant personality dimensions

As part of an attempt to standardize quality standards, the CER (Community of European Railways) and the ETF (European Transport Workers’ Federation) have drawn up a catalog of requirements listing the psychological dimensions that a suitability assessment should cover; this is more detailed and more comprehensive than the EU directive. It includes the following ability and personality characteristics:

- logical reasoning
- verbal ability (communication)
- memory
- resilience of the ability to react
- ability to react
- attention and concentration
The selection of dimensions for the SARAIL test set is based on this catalog of requirements. The ability and personality dimensions listed there have also been found in meta-analyses (Salgado, Anderson, Moscoso, Bertua, de Fruyt & Rolland, 2003) and numerous rail-specific studies (for a summary: Evans & Johnson, 2010; Johnson & Evans, 2011) to have predictive validity.

Up-to-date ideas on validity theory were taken into account when selecting the tests to measure these dimensions, in order to ensure that the psychometric quality of the tests used is as high as possible. The following section describes the process of selecting tests for the SARAIL test set. For details of the available evidence of validity, the interested reader can refer to the individual test manuals and the publications cited therein.

### 2.2 Construction of the test set

#### Guidelines on the selection of tests

When selecting tests, the most important criterion is their validity. Modern approaches to validity theory distinguish various aspects of validity (see Borsboom, Mellenbergh & van Heerden, 2004; Embretson, 1983; Kane, 2001; Messick, 1995). It should be borne in mind here that the different aspects of validity are not interchangeable; instead they combine to form a network of evidence that can be used to back up decisions taken in a particular case. Figure 2 illustrates the different types of evidence of validity that are required.

![Diagram showing the different types of evidence of validity that are required](image-url)

As part of the verification of validity it is important to check the **dimensionality** of the tests used (see Borsboom et al., 2004). This involves demonstrating that observable differences in
performance between different individuals can be unequivocally explained by differences in the particular characteristic. For example, it must be proven that a poor result on a logical reasoning test is actually attributable to poor logical reasoning ability and not to some other characteristic or ability.

The test’s construct validity must also be checked. Proving that a particular latent ability underlies the test result does not in itself guarantee that the latent ability is actually the one that is of interest. Construct validity only exists if it can be demonstrated that a test implements particular theory-led approaches. Embretson (1983) distinguishes here between construct representation and nomothetic span. **Construct representation** determines what the test score means in terms of content; it is concerned with the respondent’s cognitive processes while working the items and the extent to which these processes depend on structural features of the individual items. It must therefore be demonstrated that particular item characteristics do in fact involve the theoretically postulated cognitive processes and can therefore explain differences in the item difficulty parameters. The influence of the item characteristics and the associated cognitive processes can be empirically tested using explanatory IRT models such as the Linear Logistic Test Model (LLTM, Fischer, 1974).

As an indication of adequate construct validity, the theoretically expected correlations with other tests should also be observable. This latter aspect relates to what Embretson (1983) and Messick (1995) term the demonstration of **nomothetic span**.

In addition, the test’s **predictive criterion validity** must be checked. The purpose of psychological assessment is frequently to predict relevant human behavior. Selection and measurement of the criterion is of key importance here. In the area of rail psychology the criteria used to assess the practical relevance of the selected tests often include measures of driving behavior, the incidence of critical driving situations, supervisors’ assessments and training indices (Evans & Johnson, 2010; Johnson & Evans, 2011).

**Implementation in the construction of the SARAIL test set**
The first step involved using the requirement profiles of EU Directive 2007/59/EC, the CER (Community of European Railways) and the ETF (European Transport Workers’ Federation) to identify relevant predictors (see Section 2.1). On the basis of these criteria, tests were selected for 14 job-related ability and personality characteristics. The particular tests chosen were those that best met the requirements for evidence of validity in terms of dimensionality, construct representation and nomothetic span.

The 14 dimensions are: logical reasoning, verbal ability, memory, resilience of the ability to react, ability to react, concentration, vigilance, visual perception, eye-hand coordination, emotional stability, extraversion, conscientiousness, openness and agreeableness. On the basis of their content these can be grouped into five categories, which are illustrated in Figure 3.
2.3 Psychometric properties

Objectivity
Being normed computer tests, the tests in the SARAIL test set meet the requirements of test administrator independence, security against miscalculation and unambiguity of interpretation. In terms of test administrator independence computerized tests are in general superior to paper-and-pencil tests (Kubinger, 2006). Security against miscalculation is of course always ensured for computerized tests. All normed tests have unambiguity of interpretation. A particular percentile ranking has a defined meaning.

Reliability
Reliability describes a test’s formal exactness or precision of measurement (Lienert & Raatz, 1998). Table 1 shows the reliabilities of the individual tests in the test set.
As Table 1 shows, the reliability of the individual subtests meets the standards for psychometric tests used in suitability assessment (American Psychological Association, 1999; Häcker, Leutner & Amelang, 1998; Hornke & Winterfeld, 2004; International Test Commission, 2001; Kersting, 2008).

Validity
There is a large body of evidence supporting the validity of the tests in the SARAIL test set. Studies of their dimensionality, construct representation and nomothetic span are cited extensively in the individual test manuals. This information is supplemented by analysis of the tests’ measurement invariance across relevant sociodemographic groups. The findings quoted support the dimensionality and construct validity of the tests in the SARAIL test set. Some of the test manuals also include studies of the tests’ criterion validity in the field of transport safety.
Economy
The tests in the SARAIL test set are economical in terms of the test administrator’s time, since no time is spent giving instructions or scoring tests. In addition, if a number of test stations are available, staggered testing can be carried out.

Usefulness
A test is useful if it measures a characteristic for the assessment of which there is a practical need. A test therefore has a high degree of usefulness if it cannot be replaced by any other test (Lienert & Raatz, 1998). The practical need for the present tests arises from the need to assess cognitive abilities, stress tolerance and ability to react and to measure perception and attention functions, sensomotor functions and job-related personality characteristics when selecting people suited to be train drivers.

Reasonableness and fairness
The SARAIL test set takes around two hours to complete. The burden on candidates in terms of time and mental strain is therefore low in relation to the usefulness of the information gained; the psychometric property of reasonableness can therefore be regarded as met (Kubinger, 2006). It is often maintained that the computer is an additional stress factor in the test situation and that individuals with limited computer experience are disadvantaged. However, this does not apply to the SARAIL test set, since the special hardware for the input of responses can be easily used by anyone, irrespective of age, educational background or previous computer experience. An additional factor that contributes to the high level of fairness of the SARAIL test set is the way in which the instructions are presented. The instructions and practice phases of the individual tests are so constructed that the test itself is only presented when the task has been fully understood. Further evidence of the fairness of the tests in the SARAIL test set is yielded by studies of the measurement invariance of the tests for relevant sociodemographic groups (for details see the information on fairness in the test manuals). In summary it can be stated that the theoretical considerations and empirical findings described above in relation to the individual tests in SARAIL demonstrate that the test set is reasonable and fair.

Norming
According to Lienert und Raatz (1998) a test fulfils the quality criterion of norming if the following conditions are met: the norms are not out of date, and the populations to which the norms apply and the sample used are defined. In the context of the SARAIL test set the norm sample is a representative sample of the normal population, based on the results of censuses in Germany and Austria.
# 3 Dimensions

Table 2 provides a summary of the dimensions and tests contained in the test set. A full description of the psychometric properties and theoretical background of each test will be found in the relevant test manual. Contact us (info@schuhfried.at) – we shall be happy to send the manuals to you.

<table>
<thead>
<tr>
<th>Dimension</th>
<th>Subdimension/Form</th>
<th>Test</th>
<th>Subtest</th>
<th>Variable</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Cognitive abilities</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Logical reasoning</td>
<td>Short form</td>
<td>INSSV</td>
<td>S2/ Figural Inductive Reasoning</td>
<td>Figural Inductive Reasoning</td>
<td>15</td>
</tr>
<tr>
<td>Verbal ability</td>
<td>Short form</td>
<td>INSSV</td>
<td>S2/ Word Meaning</td>
<td>Word Meaning</td>
<td>7</td>
</tr>
<tr>
<td>Memory</td>
<td>Short-term</td>
<td>VISGED</td>
<td>S11</td>
<td>Visual short-term memory</td>
<td>13</td>
</tr>
<tr>
<td><strong>Stress tolerance and ability to react</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stress tolerance, reactive</td>
<td>-</td>
<td>DT</td>
<td>S1</td>
<td>Correct</td>
<td>6</td>
</tr>
<tr>
<td>Ability to react, simple</td>
<td>reaction speed;</td>
<td>RT</td>
<td>S3</td>
<td>Mean reaction time; mean motor time</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>motor speed</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Attention and concentration</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Concentration</td>
<td>-</td>
<td>COG</td>
<td>S11</td>
<td>Mean time correct rejection</td>
<td>10</td>
</tr>
<tr>
<td>Vigilance</td>
<td>Visual</td>
<td>WAFV</td>
<td>S2</td>
<td>Mean reaction time</td>
<td>32</td>
</tr>
<tr>
<td><strong>Sensomotor functions and perception</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Visual perception</td>
<td>-</td>
<td>LVT</td>
<td>S2</td>
<td>Score</td>
<td>13</td>
</tr>
<tr>
<td>Eye-hand coordination</td>
<td>Two-dimensional</td>
<td>2HAND</td>
<td>S3</td>
<td>Total mean duration</td>
<td>4</td>
</tr>
<tr>
<td><strong>Determinants of personality</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Emotional stability</td>
<td>Short form</td>
<td>BFSI</td>
<td>S1/ Social confidence, emotional robustness</td>
<td>Emotional stability</td>
<td>2</td>
</tr>
<tr>
<td>Extraversion</td>
<td>Short form</td>
<td>BFSI</td>
<td>S1/ Sociability, assertiveness</td>
<td>Extraversion</td>
<td>2</td>
</tr>
<tr>
<td>Conscientiousness</td>
<td>Short form</td>
<td>BFSI</td>
<td>S1/ Sense of duty, ambition</td>
<td>Conscientiousness</td>
<td>2</td>
</tr>
<tr>
<td>Openness</td>
<td>Short form</td>
<td>BFSI</td>
<td>S1/ Openness to actions, openness to ideas</td>
<td>Openness</td>
<td>2</td>
</tr>
<tr>
<td>Agreeableness</td>
<td>Short form</td>
<td>BFSI</td>
<td>S1/ Genuineness, helpfulness</td>
<td>Agreeableness</td>
<td>2</td>
</tr>
<tr>
<td><strong>Total length</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>120</td>
</tr>
</tbody>
</table>

## 3.1 Cognitive abilities

**Logical reasoning**
Logical reasoning is also termed fluid intelligence; it involves the ability to recognize regularities, understand the implications of statements and draw logical conclusions (Arendasy & Sommer, 2012; Carroll, 1993; Horn & Noll, 1997). Fluid intelligence is largely innate or inherited; the environment has little impact on it. It is therefore also termed cognitive potential and is a very good predictor of long-term occupational success.

Test:
Logical reasoning is measured using the Figural Inductive Reasoning subtest from the INSSV (Intelligence Structure Battery – Short Form, S2). In this test the candidate is shown a 3x3 matrix; eight of the nine fields contain figures that are arranged in accordance with various rules. The candidate must identify these rules and apply them in order to fill in the empty ninth field. The short form used in this test set has a time limit of 15 minutes.

Verbal ability
Verbal ability describes an aspect of a person’s crystallized intelligence. Crystallized intelligence includes all skills that are learned in the course of life or determined by the environment. In the SARAIL test set verbal ability is viewed as the extent of a person’s vocabulary (Carroll, 1993; Horn & Noll, 1997). This dimension does not assess an individual’s knowledge of grammar or his ability to express himself. People with good verbal ability are better at communicating with people from different occupational backgrounds and applying their specialist knowledge appropriately in concrete work situations.

Test:
Verbal ability is assessed by the Word meaning subtest from the INSSV (Intelligence Structure Battery – Short Form, S2). In this test the candidate is presented with a succession of terms and must select from the four alternatives offered the term that is most similar in meaning to the starting term. The short form used in this test set has a time limit of 7 minutes.

Memory
Memory is the ability to retain, organize and later retrieve perceived information. A distinction is made between short-term and long-term memory (see Carroll, 1993; Horn & Noll, 1997). In SARAIL only short-term memory is tested, and this is done using visual material. The test set therefore tests the ability to retain visual information in the short term and reproduce it correctly.

Test:
Short-term memory is measured by the VISGED test (Visual Memory Test, S11). In this test a street map is shown briefly on the screen. The map is marked with different symbols (e.g. a cross for the hospital, a book for the library, etc.). The nature and position of the symbols must be memorized and later remembered correctly.

3.2 Stress tolerance and ability to react

Stress tolerance, reactive
Reactive stress tolerance is the individual’s ability to react quickly and accurately when under stress. To place the candidate in a stressful situation, stimulus presentation is adapted to his or her ability level.

Test:
Reactive stress tolerance is measured by the DT (Determination Test, S1). The test uses five optical stimuli of different colors, two different acoustic stimuli and two signals for the foot-operated keys. To all of these the candidate must react as quickly as possible by pressing the corresponding button on the Response Panel or the corresponding foot pedal. The program adapts to the candidate’s ability level by varying the speed of stimulus presentation.
Ability to react, simple
Ability to react involves the ability to respond to one or more stimuli as quickly and accurately as possible. In the case of simple reaction ability, the motor action is triggered by a simple signal. This signal is followed by a pre-defined sequence of movements. The signal may be delivered acoustically, visually or kinesthetically. Adequate reaction speed is essential when in charge of a vehicle, since when an incident occurs suddenly even milliseconds may count.

A distinction is made between reaction speed and motor speed. Reaction time is the time that elapses between stimulus presentation and the start of the mechanical response movement (i.e. the point at which the finger leaves the rest button). Motor time is the time that elapses between the moment the finger leaves the rest button and the time the reaction button is pressed. This score provides information about the respondent’s speed of movement.

Test:
Simple reaction ability is measured in the SARAIL test set by the RT (Reaction Test, S3). In this test the candidate must press a particular button as quickly as possible as soon as the required stimulus appears. The Response Panel is needed for administration of this test. It is possible to distinguish between “reaction speed” and “motor speed”.

3.3 Attention and concentration

Concentration
Concentration is regarded in this context as a special function of attention, namely selective attention. When in charge of a vehicle, this is called on every time that a stimulus needs to be shielded from other stimuli and concentration needs to be shifted from one stimulus to another.

Test:
Concentration is measured in the SARAIL test set by the COG (Cognitrone, S11). Unlike traditional tests of concentration, the Cognitrone contains item material of varying complexity. The candidate’s task is to compare a figure with four other figures and to state as quickly as possible whether one of the four figures is identical to the reference figure.

Vigilance
Vigilance is the ability to focus attention persistently over lengthy periods of time on one or more sources of information, in order to detect and respond to small changes in the information received (e.g. Zomeren & Brouwer, 1994). The relevant stimuli typically occur only at very irregular intervals and at very low frequency among a large number of irrelevant stimuli. Vigilance can be measured both visually and aurally.

Test:
In the SARAIL test set vigilance is measured by the WAFV test (Perception and Attention Functions: Vigilance, S2). In this test the candidate is presented with visual or auditory stimuli that occasionally diminish somewhat in intensity. The task is to react in these occasional cases, which constitute about 5% of the stimuli.

3.4 Sensomotor functions and perception

Eye-hand coordination, two-dimensional
Eye-hand coordination tests how well someone can convert visual information into hand movements – in other words, how well they can coordinate eye and hand in making fine, small-scale movements.

Test:
In the SARAIL test set eye-hand coordination is measured by the **2HAND** test (Two-hand Coordination, S2). This assesses both sensomotor coordination between eye and hand and coordination between left and right hand. The respondent’s task is to move a red dot along a prescribed track as fast as possible, using two twist knobs or joysticks. The track consists of three sections that make different demands on the coordination of left and right hand.

**Visual perception**
Visual perception involves the reception and processing of visual stimuli. Since the beginning of the 1960s line maze tests have been used to check visual orientation when testing fitness to drive or studying perception in driving situations.

**Test:**
In the SARAIL test set visual perception is measured by the **LVT** (Visual Pursuit Test, S2). The respondent is presented with an array of lines and must as quickly as possible find the end of a specified line. Tracing the lines with a finger is not permitted. There is no time limit on the test.

### 3.5 Determinants of personality

**Emotional stability**
Emotional stability describes how people tend to deal with their emotions and with potential stresses. Individuals with a high level of emotional stability see themselves as emotionally stable, calm, even-tempered, reflective and optimistic; even in stress situations they do not readily lose their composure. By contrast, individuals with low emotional stability regard themselves as anxious; they more frequently report negative feelings, see themselves as readily becoming irritated and losing control and feel insecure in social interaction with others.

**Test:**
Emotional stability is measured by the **BFSI** (Big Five Structure Inventory, S1). The BFSI is a state-of-the-art test for comprehensive assessment of the personality structure, based on the Big Five model. In the SARAIL test set emotional stability is assessed by means of the two facets of social confidence and emotional robustness. Using a four-point answer scale, respondents indicate the extent to which various adjectives or statements apply to them. In interpreting an individual’s results it should be borne in mind that the information obtained is a self-assessment. The test scores thus reflect the self-image that the individual has of himself or that he wishes to portray.

**Extraversion**
Extraversion describes individual tendencies in interpersonal behavior. Individuals who score highly on extraversion describe themselves as warm, sociable, outgoing, good-humored, assertive, self-confident, energetic and adventurousness. Individuals with lower levels of extraversion see themselves by contrast as somewhat reticent and safety-conscious and having little interest in asserting themselves. They prefer to be alone and independent or else to interact with relatively small groups.

**Test:**
In the SARAIL test set extraversion is assessed via the facets of sociability and assertiveness in the **BFSI** (Big Five Structure Inventory, S1). Using a four-point answer scale, respondents indicate the extent to which various adjectives or statements apply to them. In interpreting an individual’s results it should be borne in mind that the information obtained is a self-assessment. The test scores thus reflect the self-image that the individual has of himself or that he wishes to portray.
Conscientiousness
Conscientiousness describes personal tendencies in someone’s work behavior. People with a high level of conscientiousness describe themselves as competent, organized, orderly, dutiful, focused, disciplined and ambitious. By contrast, individuals with a lower level of conscientiousness see themselves as not very focused or disciplined at work, less precise and not particularly ambitious.

Test:
In the SARAIL test set conscientiousness is assessed via the facets of sense of duty and ambition in the BFSI (Big Five Structure Inventory, S1). Using a four-point answer scale, respondents indicate the extent to which various adjectives or statements apply to them. In interpreting an individual’s results it should be borne in mind that the information obtained is a self-assessment. The test scores thus reflect the self-image that the individual has of himself or that he wishes to portray.

Openness
Openness describes individual tendencies in dealing with new experiences, impressions, ideas and values. Individuals with a high level of openness often state that they have a lively imagination, are interested in a wide range of things and are particularly hungry for knowledge. They also describe themselves as tolerant and open to other views and values. By contrast, individuals who are less open are more inclined towards conventional attitudes and activities. They prefer the familiar to the new and describe themselves as having little interest in things in general or in the acquisition of new knowledge.

Test:
Openness is measured by the BFSI (Big Five Structure Inventory, S1). The SARAIL test set uses the two facets of openness to actions and openness to ideas. Using a four-point answer scale, respondents indicate the extent to which various adjectives or statements apply to them. In interpreting an individual’s results it should be borne in mind that the information obtained is a self-assessment. The test scores thus reflect the self-image that the individual has of himself or that he wishes to portray.

Agreeableness
Agreeableness describes individual tendencies in people’s way of handling interpersonal relationships. People with a high level of agreeableness see themselves as obliging, understanding, considerate and helpful. They are fairly modest, trust other people and often act cooperatively. People with a low level of agreeableness describe themselves as more self-referential and not particularly accommodating. In addition they admit to being less helpful and cooperative, and they are often mistrustful of other people.

Test:
In the SARAIL test set agreeableness is measured via the facets of genuineness and helpfulness in the BFSI (Big Five Structure Inventory, S1). Using a four-point answer scale, respondents indicate the extent to which various adjectives or statements apply to them. In interpreting an individual’s results it should be borne in mind that the information obtained is a self-assessment. The test scores thus reflect the self-image that the individual has of himself or that he wishes to portray.
4 SCORING AND INTERPRETATION

General information on the use of the scoring options can be found in the Vienna Test System help file. This section contains scoring and information guidelines specific to the SARAIL test set.

In order to maximize the validity of the assessment of a candidate’s suitability, all available information about the individual in question should be incorporated into the assessment. The primary information is the normed scores on the individual tests. However, when assessing suitability the test results should always be interpreted in the light of the specific requirements profile. Regional or company-specific rules should also be borne in mind.

4.1 Results on each dimension

The individual test results are presented in the SARAIL test set in the form of raw scores as well as parameters and percentile ranks. The percentile rank indicates what percentage of a particular comparison group achieved the same or a lower score on the ability or personality characteristic in question. The comparison group is a representative sample of the general population. For example, a percentile rank of PR=76 means that 76% of respondents in the representative norm sample obtain this score or a lower one on this characteristic, and 24% obtain a higher score.

- A percentile rank of 24 or less indicates a below-average level of the ability or personality characteristic in question by comparison with the representative norm sample.
- A percentile rank of 25 – 75 reflects an average level of the ability or personality characteristic in question by comparison with the representative norm sample.
- A percentile rank of 76 or more indicates an above-average level of the ability or personality characteristic in question by comparison with the representative norm sample.

4.2 Standardization

Table 3 gives an overview of the representative norms that have been used for the SARAIL test set.
An individual written report can be produced for each candidate. The report contains a profile diagram of the test results. The report can be used as the basis for an interview or for feedback of results to candidates.

4.4 Case studies

Case study 1
Mr. Smith is applying for a job as a train driver with a regional rail company. As part of the application process, he completes the SARAIL test set in a two-hour test session.

Table 4 shows his percentile ranks on the individual dimensions. Above-average test scores are shown in green, average test scores are shown in black and below-average scores are printed in red.
### Table 4: Case study 1

<table>
<thead>
<tr>
<th>Dimensions</th>
<th>PR</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Cognitive abilities</strong></td>
<td></td>
</tr>
<tr>
<td>Logical reasoning (INSSV)</td>
<td>77</td>
</tr>
<tr>
<td>Verbal ability INSSV)</td>
<td>65</td>
</tr>
<tr>
<td>Memory (VISGED)</td>
<td>72</td>
</tr>
<tr>
<td><strong>Stress tolerance and ability to react</strong></td>
<td></td>
</tr>
<tr>
<td>Ability to react (RT)</td>
<td>87</td>
</tr>
<tr>
<td>Stress tolerance, reactive (DT)</td>
<td>73</td>
</tr>
<tr>
<td><strong>Attention and concentration</strong></td>
<td></td>
</tr>
<tr>
<td>Concentration (COG)</td>
<td>75</td>
</tr>
<tr>
<td>Vigilance (WAFV)</td>
<td>82</td>
</tr>
<tr>
<td><strong>Sensomotor functions and visual perception</strong></td>
<td></td>
</tr>
<tr>
<td>Eye-hand coordination (2HAND)</td>
<td>92</td>
</tr>
<tr>
<td>Visual perception (LVT)</td>
<td>95</td>
</tr>
<tr>
<td><strong>Determinants of personality</strong></td>
<td></td>
</tr>
<tr>
<td>Emotional stability (BFSI)</td>
<td>74</td>
</tr>
<tr>
<td>Extraversion (BFSI)</td>
<td>45</td>
</tr>
<tr>
<td>Conscientiousness (BFSI)</td>
<td>72</td>
</tr>
<tr>
<td>Openness (BFSI)</td>
<td>65</td>
</tr>
<tr>
<td>Agreeableness (BFSI)</td>
<td>74</td>
</tr>
</tbody>
</table>

By comparison with the norm sample Mr. Smith has above-average scores on the dimensions of logical reasoning (INSSV), ability to react (RT), vigilance (VIGIL), eye-hand coordination (2HAND) and visual perception (LVT). His other scores on both ability and personality dimensions are all average. Mr. Smith therefore has satisfactory scores in all the areas that were tested; he can therefore be recommended for inclusion in the next stage of the application process.

### Case study 2

Mr. Jones is applying to the public transport authority for retraining as an underground train driver. As part of the internal application process, he completes the SARAIL test set in a two-hour test session.

Table 5 shows his percentile ranks on the individual dimensions. Above-average test scores are shown in green, average test scores are shown in black and below-average scores are printed in red.
### Table 5: Case study 2

<table>
<thead>
<tr>
<th>Dimensions</th>
<th>PR</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Cognitive abilities</strong></td>
<td></td>
</tr>
<tr>
<td>Logical reasoning (INSSV)</td>
<td>35</td>
</tr>
<tr>
<td>Verbal ability INSSV)</td>
<td>45</td>
</tr>
<tr>
<td>Memory (VISGED)</td>
<td>75</td>
</tr>
<tr>
<td><strong>Stress tolerance and ability to react</strong></td>
<td></td>
</tr>
<tr>
<td>Ability to react (RT)</td>
<td>12</td>
</tr>
<tr>
<td>Stress tolerance, reactive (DT)</td>
<td>4</td>
</tr>
<tr>
<td><strong>Attention and concentration</strong></td>
<td></td>
</tr>
<tr>
<td>Concentration (COG)</td>
<td>22</td>
</tr>
<tr>
<td>Vigilance (WAFV)</td>
<td>20</td>
</tr>
<tr>
<td><strong>Sensomotor functions and visual perception</strong></td>
<td></td>
</tr>
<tr>
<td>Eye-hand coordination (2HAND)</td>
<td>84</td>
</tr>
<tr>
<td>Visual perception (LVT)</td>
<td>27</td>
</tr>
<tr>
<td><strong>Determinants of personality</strong></td>
<td></td>
</tr>
<tr>
<td>Emotional stability (BFSI)</td>
<td>31</td>
</tr>
<tr>
<td>Extraversion (BFSI)</td>
<td>70</td>
</tr>
<tr>
<td>Conscientiousness (BFSI)</td>
<td>84</td>
</tr>
<tr>
<td>Openness (BFSI)</td>
<td>74</td>
</tr>
<tr>
<td>Agreeableness (BFSI)</td>
<td>64</td>
</tr>
</tbody>
</table>

Mr. Jones obtains an above-average score in the area of eye-hand coordination (2HAND). By comparison with the norm sample he is therefore very good at coordinating eye and hand when making fine, small-scale movements. On the personality side he describes himself as particularly competent, organized, orderly, conscientious, striving, disciplined and ambitious (BFSI – conscientiousness). However, his ability to react is poor (R), as is his reactive stress tolerance (DT). This means that Mr. Jones is not sufficiently good at reacting quickly and appropriately when under stress. In addition, he has deficits in concentration (COG) and vigilance (WAFV). Overall, therefore, it must be assumed that his attention and reaction skills are inadequate. On account of his poor attention and reaction skills, Mr. Jones will not be included in the next stage of the application process.
5 REFERENCES


6 INDEX OF TABLES AND ILLUSTRATIONS

Table 1: Reliability and length of the tests in SARAIL ......................................................... 8
Table 2: Summary of the dimensions in the SARAIL test ..................................................... 10
Table 3: Overview of the standardization of the SARAIL test set ...................................... 16
Table 4: Case study 1 ........................................................................................................... 17
Table 5: Case study 2 ........................................................................................................... 18

Figure 1: Overview – manuals .............................................................................................. 3
Figure 2: Diagram showing the different types of evidence of validity that are required ... 5