



“Technical Assistance for the SA-EU

Dialogue Facility”

(EuropeAid/137708/DH/SER/ZA)

Report:

*Desktop Research of Non-Psychometric Career Assessment Guidelines in
the European Union Member States*

Dialogue project:

Senior Non-Key Expert

Dr. John McCarthy LL.D (Hon.), C.Psychol., AFBPS



A project implemented by

DAI human dynamics
Human & Organisational Dynamics Limited, KG

	Project Principal	Senior Non-key Expert (NKE)
Name:	SA-EU Dialogue Facility	John McCarthy
Address	Unit 5, Parklands Office Park 223 Bronkhorst Street Brooklyn, Pretoria SOUTH AFRICA	161 B AVENUE SAINTE MARGUERITE, 06200 NICE, FRANCE
Contact	Ms Vuyelwa Masagwana	John McCarthy
Telephone	+27 (0)82 339 2131	+33 671779263
E-mail	Teamleader@dialoguefacility.org	jmcguidance@gmail.com

Executive summary.....	5
Introduction.....	6
CHAPTER 1 KEY CONCEPTS AND METHODS FOR MENTAL MEASUREMENTS.....	6
CHAPTER 2 NON-PSYCHOMETRIC TESTS.....	11
CHAPTER 3 ONLINE TESTING CONSIDERATIONS.....	15
CHAPTER 4 THE KNOWLEDGE AND COMPETENCE OF CAREER PRACTITIONERS IN ASSESSMENT.....	21
CHAPTER 5 ETHICS OF TESTING FOR CAREER PRACTITIONERS.....	25
CHAPTER 6 REVIEW OF FREE ONLINE PUBLICLY FUNDED CAREER ASSESSMENT TESTS ON GOVERNMENT SUPPORTED WEBSITES IN EUROPEAN COUNTRIES AND NATIONS ¹	26
CHAPTER 7 TOWARDS A SET OF GUIDELINES FOR THE DEVELOPMENT OF NON-PSYCHOMETRIC CAREER LEARNING TESTS.....	31
CHAPTER 8 ADDITIONAL RECOMMENDATIONS FOR SOUTH AFRICA.....	37
REFERENCES.....	37
ANNEX ONE: EUROPEAN UNION AND NATIONS -TABLE OF FREE ONLINE TESTS PROVIDED ON GOVERNMENT SUPPORTED WEBSITES.....	40

¹ Please refer to Methodological Note in Annex Two

ANNEX TWO: METHODOLOGICAL NOTE ON REVIEW
OF EUROPEAN UNION WEBSITE
TESTS.....46

ANNEX THREE: LIST OF PERSONS COMMUNICATED WITH FROM WHOM
ASSISTANCE WAS RECEIVED FOR THIS
REPORT..... 48

EXECUTIVE SUMMARY

This study aimed to find the basis for draft proposals for guidelines for non-psychometric tests drawing on an international literature review and, on the practice and experience of European countries in the context of South Africa-European Union Strategic Partnership Dialogue Facility.

The literature review revealed the existence of internationally recognized standards for psychometric tests, an internationally recognized review model for psychometric tests, and internationally recognized guidelines for internet/online and computerized testing. These standards support best practice in psychometric test development and usage, but they are not regulatory, not legal requirements, and not enforced by the sponsoring professional organisations.

With regard to non-psychometric tests, no such standards or guidelines exist. Some proposals for guidelines have been made by some researchers but in general the field of non-psychometric testing has been largely ignored by professional associations and researchers. Competence in the use of career assessments has been identified as a key competence for career practitioners in an EU-wide study, and examples can be found at national level of government training specification requirements for this competence for career practitioners. At national level, examples can be found of professional association ethical guidelines for test usage by career practitioners along with some guidelines for career practitioners for online testing.

The review of free publicly funded online career assessment tests in 22 European countries and nations shows that the majority of governments provide tests of a non-psychometric nature. Their main function is to support self-appraisal and career exploration. No guidelines or standards exist for the development of such tests, but theoretical and career expert, and more recently, multidisciplinary team input, are evident in their development. Online tests are most commonly tests of occupational interests while some countries additionally provide tests of abilities, skills, decision-making and career planning. More than half of the countries have separate tests for youth and for adults. Many tests allow test-taker results to be matched with education and training pathways and with occupations. Funding for youth tests comes mainly but not exclusively from the ministry of education while funding for tests for adults comes mainly but not exclusively from the ministry of labour. While online tests are standalone tools for self-appraisal and career exploration, countries offer varying degrees of support to test-takers. Online registration allows for test usage and test-taker evaluation. There is an increasing recognition that online test design and development requires the input of multidisciplinary teams and the use of Gameful Design principles.

Based on the international literature review and on the review of experience and practice in 22 European countries and nations, draft proposals for guidelines for non-psychometric tests are outlined according to eight categories: validity, reliability, fairness, design and development, scoring and administration, documentation, test-user responsibilities, online and computerized testing. Some additional recommendations are made.

INTRODUCTION TO THE REPORT

The context of this report is the learning environment in South Africa that enables people of all ages -individuals, families, communities, ethnic groups - to make informed choices of occupation and work opportunities in the labour market and of the learning pathways (education, training) to these opportunities. People (individuals, families, groups) differ in their capacity to visualize their futures, to identify, source, and interpret information about themselves, family and societal expectations, and learning and work opportunities, and to apply this knowledge in decision-making (ELGPN, 2015). People also differ in the pace of their emotional and social development that similarly influences their decision-making. Given the range of human differences, many people will benefit from some form of career practitioner support for decision-making, with a significant minority requiring a lot of support. Most theories of career development agree that knowledge of oneself is a key ingredient in any career decision making (Osborn and Zunker, 2016). Self-knowledge, acquired through non-formal and formal learning, includes knowledge about one's interests, values, beliefs, skills, identity, self-esteem, personal qualities, needs, circumstances etc. Psychometric and non-psychometric career assessment tests are a recognized means of increasing self-knowledge. This report will address some of the key issues in the construction, development, and usage of such tests. It will further examine how a number of European countries have addressed such issues in the provision of publicly funded free online tests.

CHAPTER 1 KEY CONCEPTS AND METHODS FOR MENTAL MEASUREMENTS

Tests are formalized ways of observing and/or measuring different aspects of mental functioning e.g. abilities, achievements, interests, values, personality, that help individuals and professionals to have a more precise and complete understanding of their mental processes and capacities. They are used in a variety of settings such as education, health, employment, sports, and social affairs to enable self-knowledge, other-knowledge, diagnosis, exploration, decision-making, differentiation, comparison, selection, prediction, and prescription. Tests can have written/printed formats or be computerized or online and can be taken with or without administrative supervision or support.

Tests can be standardized and norm-referenced (psychometric) or non-standardised (non-psychometric). Persons who administer psychometric tests (known as 'test users'²) require formal specialized training in psychometric theory, in the theory (if applicable) underlying a test, in the practice of using these tests, and in interpreting the test scores derived. Similar and different training is useful for persons who develop and administer non-psychometric tests e.g. formative subject tests designed and administered by teachers. Non-psychometric tests are also to be found in books, magazines, and other printed materials, in computer programmes, and on online websites. Many of these are self-administered by the general public seeking some form of

² In this report, the term 'test user' refers to persons who select, administer, and score tests and interpret test scores. Persons who undertake tests are referred to as "test takers".

self-knowledge. A common concern for all test users and test takers, including the general public, is how to select, administer, and score tests, and how to interpret their results.

This chapter explores issues in mental measurements, how to manage measurement errors, test fairness, and a methodology for reviewing psychometric tests. This knowledge will help differentiate psychometric tests from non-psychometric tests.

1.1 Measuring mental constructs

Psychometrics refers to the quantification and measurement of mental processes that we use, for example, to remember, to learn, to reason, to believe, to understand, and to describe different aspects of our personality. Such measurements are useful for self-understanding, for understanding individual differences, for peer and group comparisons, for career and other decision-making, and for predicting future behaviour. We are constantly processing information, such as in reading this text, and constructing and reconstructing how we understand our world through this process of cognition. Simultaneously, this new information informs our knowledge, attitudes, and behaviours. Hence, psychometrics refers to the measurement of cognition in different manifestations e.g. ability, aptitude, beliefs, attitudes, values, preferences, knowledge, skills, behaviours etc.

Psychometric methods concern the application of statistical designs to the construction, development, interpretation, and usage of psychological tests and measures. The start point of test development is the definition of a construct e.g. extroversion, and the transformation of this construct into a set of variables e.g. test items or performance exercises. An individual's responses to these test items are quantified, and inferences are drawn concerning the individual's performance relating to the construct. Such inferences can also be used for inter-individual or inter-group comparisons and/or for prediction. However, each step above is fraught with risk! Who defines the construct? How well is the construct defined? How well do the variables represent the construct? How representative of the construct are the particular set of variables? What can really be inferred from the quantification of the test responses and from the manner of that quantification? Who can administer such tests? Who can make such inferences? Can such tests be self-administered and self-interpreted? What standards or guidelines exist for test construction and usage in order to reduce these risks? How are people trained to use tests? Are there ethical and legal issues in testing?

1.2 Managing sources of error in mental measurement

There are many possible sources of error when we apply quantification to measuring mental constructs. Some of these are random by nature e.g. the physiological or motivational condition of the individual at the time of taking the test; or the conditions of test taking. Other sources of error are systematic e.g. a cultural or age response bias to particular test items that could skew the result achieved in a test or the existence of an imperfect representation of a construct in test items leading to a misleading interpretation of a test score.

To obviate or minimize the effects of such errors, mental measurement theorists in education and psychology have advanced the concepts of the validity and reliability of test scores. Validity refers to 'the degree to which interpreting and using test scores for a particular decision is supported by evidence and theory' (Osterlind, 2010). The industry standards for test

development, *Standards for Educational and Psychological Testing* (AERA, 2014)³, hold that the evaluation of validity is the joint responsibility of the test developer and the test user (person who administers and interprets the test): “The test developer is responsible for furnishing relevant evidence and a rationale in support of the intended use. The test user is ultimately responsible for evaluating the evidence in the particular setting in which the test is to be used.” (p. 13). The *Standards* cite five sources of evidence for test score validity evaluation: test content, test response processes, the internal structure of a test, relationship of test scores to variables external to the test, and the consequences (intended, unintended) of testing. In all, 25 standards are prescribed for test score validity evaluation. As the authors point out, these are not a checklist; each type of evidence is not required for all tests. However, evidence is needed for propositions that underlie each test score interpretation. The validation of test score interpretation is an ongoing process; it is not a once off event. The amount and types of evidence required to make a provisional validity evaluation varies; the higher the stakes in the use and interpretation of test scores, the higher the standards of evidence required. As the *Standards* state, ‘validity is the most fundamental consideration in developing and evaluating tests’ (p. 11).

The reliability of test scores refers to precision in measurement, for example, how well do a sample of test items represent the universe of possible test items for a particular construct? How consistent are one’s test scores across replications of the same test and what is the impact of this on test score interpretation? How consistent are test scores across different population groups? How do different sources of error contribute to overall measurement error? The *Standards* describe eight clusters (and 21 standards) of factors to consider and procedures for establishing the reliability of test scores: replications of testing procedures; evaluating reliability/precision; reliability/generalizability coefficients; factors affecting reliability/precision; standards errors of measurement; decision consistency; reliability/precision of group means; and documenting reliability/precision. There are several different potential sources of errors in test scores and accordingly different ways to estimate test score reliability. As in the case of validity, the standards are not a check list; and the amount and type of evidence required to evaluate the reliability of test scores depends on the significance or importance of the decisions made on the basis of such scores. Test developers and distributors have primary responsibility for obtaining and reporting evidence for reliability/precision.

1.3 The fairness effect on test performance and score interpretation

The *Standards* draw attention to a range of factors of a non-psychometric nature that can impact on test performance and test score interpretation. These concern different aspects of the issue of fairness: (i) fairness in treatment of test takers during the testing process; (ii) fairness as a lack of measurement bias; (iii) fairness in access to the construct as measured; and (iv) fairness in the validity of individual test score interpretations for intended uses. The first of these is ensured by the use of standardized test administration procedures e.g. common directions for test takers, specified room arrangements/conditions, specified time limits, and for some tests, qualified test administrators. The second refers to measurement bias (item and test functioning) or bias in predictive relations by test takers based on their membership of a subgroup in society. Related to

³ Referred to as the *Standards* hence forward throughout this text.

this is the differential meaning attributed by individuals and/or groups within the intended population of test takers to the particular construct being measured in a test. Such sources of measurement bias can be minimized at the test development stage through small scale try outs, interviews, or focus groups with relevant subgroups. The third, access to the construct, refers to measurement bias in test performance arising from individual test taker's characteristics e.g. age, language proficiency, gender, disability, ethnicity, or any combination of these. If one or any combination of these characteristics are obstacles to optimal test performance of a test taker, this has implications for the interpretation of test scores. The fourth, the validity of individual test score interpretations, draws attention to the importance of taking into account the individual characteristics of the test taker and of the test taking situation when making interpretations of test scores. The *Standards* provide twenty propositions/standards that address the issue of fairness.

1.4 Other relevant considerations in psychometric test development and score interpretation

The *Standards* provide additional sets of standards that cover

- test design and development
- scores, scales, norms, score linking, and cut scores
- test administration, scoring, reporting, and interpretation
- supporting documentation for tests
- the rights and responsibilities of test takers
- the rights and responsibilities of test users.

They also offer advice on specific applications:

- psychological testing and assessment
- workplace testing and credentialing
- educational testing and assessment
- uses of tests for program evaluation, policy studies, and accountability.

The *Standards* are a rich and authoritative resource for psychological and educational test developers and publishers to guide them to produce well-constructed and well-interpreted tests. They provide a basis for evaluating such tests and testing practices. They are also a source of guidance on test selection and test score interpretation for persons who use tests as part of their work e.g. psychologists. They are not, however, a statement of legal requirements or regulations, nor are they enforceable by the organisations that contributed to their development and who sponsor and publish them. They have been recognized as professional standards for test development and practice by regulatory authorities and courts in the USA and thus merit consideration by test developers, publishers, and users in the USA. But as the authors point out, the acceptability of a test or testing practice does not rest on the literal satisfaction of every standard: "Specific circumstances affect the importance of individual standards, and individual

standards should not be considered in isolation” (p. 7). The *Standards* are regularly revised in light of developments in the field.

1.5 How to evaluate psychometric tests: a psychological test review model

The European Federation of Psychologists Associations (EFPA) has developed a Test Review Model (TRM, latest revision 2013) to enable the rigorous evaluation of psychological assessment tests, scales and questionnaires used in the fields of work, education, health and other contexts. For its sources, the TRM has mainly drawn on test evaluation review mechanisms developed by some of EFPA’s member organizations: the British Psychological Society (BPS), the Spanish Psychological Association, and the Dutch Association of Psychologists. The TRM is divided into three main parts: (i) description of the instrument (all the features of the test are described in detail); (ii) evaluation of the instrument (the fundamental properties of the test - test materials, norms, reliability, validity, and computer generated reports, including a global final evaluation); and (iii) bibliography (the references used in the review).

The description section examines the general description of the test provided by the developers/publishers, the classification (e.g. content domain, intended use and target population, demands on the test taker etc.), measurement and scoring (e.g. scoring procedure, scales used), computer generated reports (e.g. complexity, sensitivity to context), and supply conditions and costs (e.g. start up and recurrent costs, test user qualifications required for the supply of the tests).

The instrument evaluation section pays attention to quality of the explanation of the rationale, the adequacy of documentation available to the user, the quality of procedural instructions provided for the user, the quality of the test materials (paper-and-pencil tests, Computer Based Tests (CBT) or Web Based Tests (WBT), norms (norm-referenced interpretation, domain and criterion referenced interpretation), reliability, validity (construct; criterion-predictive, concurrent, postdictive; and overall validity), and the quality of computer generated reports (e.g. reliability, validity, fairness, acceptability).

The test review is intended to be undertaken by two independent reviewers, in a peer review process. A consulting editor oversees the reviews and may call in a third reviewer if significant discrepancies between the two reviews are found. The results of a test’s evaluation are normally available for all professionals and users (either paid or for free), for example, on the website of a national psychological association. The EFPA notes that all European countries do not have their own review procedures, and therefore local issues may necessitate changes in the TRM or in the review procedures when countries start to use it. Local adaptations are possible to guarantee a better fit with local needs.

Reviews based on the TRM enable test users and professionals to evaluate all the technical information about a test in order to enhance their usage, to improve tests and testing, and to inform test selection.

1.6 Summary

Measuring mental constructs through test development is a complex, technical, and lengthy endeavour. There are many possible sources of error, systematic and/or random, that can arise in

applying quantification to measuring mental constructs and when interpreting and using test scores in decision-making. It is important therefore to establish the validity and reliability of test scores and to assess test fairness factors relating to test performance and test score interpretation.

Psychometric tests are used for diagnosis, prediction, decision-making, differentiation, comparison, selection, prediction, and prescription. These tests are normally administered, scored and interpreted by persons who have acquired the knowledge, skills, training and qualifications for these tasks. The tests are usually administered in standard conditions with the presence and assistance of a professional. The higher the stakes in the use and interpretation of test scores e.g. diagnosis, selection, prediction, prescription, the higher the standards of evidence required from test developers and publishers for the interpretation of the test scores. Hence, the need for standards for test development.

The *Standards* for psychometric tests, developed in the USA, provide a basis for assisting psychometric test development and for evaluating and reviewing psychometric tests and testing practices. They supply guidance for test selection and test score interpretation to persons who use tests as part of their work e.g. psychologists. They are not, however, a statement of legal requirements, nor regulatory, nor are they enforced or enforceable by the organisations that contributed to their development and who sponsor and publish them. They have, however, been recognized as professional standards for test development and practice by some regulatory authorities and courts in the USA and thus merit consideration by test developers, publishers, and users in the USA (as similar national standards have not been developed by other countries, the USA *Standards* are often used as an international reference point). The acceptability of a test or testing practice does not rest on the literal satisfaction of every standard. The context of testing, including the purpose of test usage, may attach more importance to some standards than to others. Test developers and distributors have primary responsibility for obtaining and disseminating evidence for test score reliability, validity and fairness.

CHAPTER 2 NON-PSYCHOMETRIC TESTS

Non-psychometric tests refer to non-standardised tools developed to help clients to (i) become self-aware of relevant aspects of their personality and identity or of their needs; (ii) articulate these, and (iii) express concerns, thoughts or feelings. These tools may include checklists of interests, values, beliefs, skills, hobbies, worries, coping strategies, etc. that may provide non-standardised self-estimates. They can have different formats e.g., cards, drawings, genograms, workbooks, games, role plays, simulations, and work sampling. Non-psychometric tests are often self-administered and self-interpreted, without the support of a professional or practitioner. In the context of career learning, the use of psychometric and non-psychometric tests is perceived as aiding the development of self-knowledge and self-understanding and as learning tools that support an individual's career exploration activities and career decision-making.

2.1 Defining non-psychometric tests

In the field of career guidance, non-psychometric tests have been termed ‘non-standardised’ (Osborn and Zunker, 2016), and ‘nonstatistical’ and ‘qualitative’ assessments (Goldman, 1992) or ‘informal’ (Niles and Harris Bowsbey, 2021; Okocha, 1998). McMahon et al. (2019) noted the amorphous nature and lack of specificity of such descriptions and proposed as a definition “a structured qualitative instrument, technique or process that facilitates participant reflection”. They observed that no comprehensive review of qualitative career assessments exists that could provide a synthesis of the topic in contemporary career development theories and practice. The respective roles and values of quantitative and qualitative assessments in career learning have also been framed in terms of theoretical approaches to career guidance: quantitative is signaled as a product or concomitant of logical positivist approach, applying psychological objectivity; qualitative is viewed as complementary and integral to a constructivist narrative approach with an emphasis on subjective experience (McMahon et al., 2019). In the latter review, the authors noted that almost two-thirds of the articles on qualitative career assessments in eight widely known English language career development journals from 1993 to 2017 made some reference to the theory underlying the qualitative instruments.

Non-psychometric assessments are sometimes created by a psychologist or a counsellor or a teacher/trainer out of professional necessity (existing psychometric tools are not fit for purpose) or economic necessity (funds not available to develop or to purchase relevant psychometric tools). They are to be found in printed publications on self-help in the fields of mental health, career guidance, job search among others, and online in publicly and privately funded websites dedicated to such fields. In general, they are self-guided materials for experiential learning, useful both to the individual in terms of self-help and to professionals to enable them to have a better understanding of their clients. Niles and Harris-Bowsbey (2021) state that informal assessments may be less anxiety producing for clients while Goldman (1992) reports that they are particularly useful for group work learning. They offer more opportunity for self-expression to the client than psychometric tests, thus encouraging the active engagement of the participant. They are flexible and adaptable for use with diverse population groups e.g. life experience, age, ethnicity, culture, learning disability and other types of disability, and learning style. They can complement (though not act as a substitute for) client information gathered through psychometric tests and other means and can be used to stimulate discussions or to raise questions.

What are the limitations of non-psychometric tests? Of their nature, they lack the statistical types of data that accompany psychometric tests such as norms, expectancy tables, interpretative charts, and statistical estimates of reliability and validity; they lack scientific rigour (Niles and Harris-Bowsbey, 2021). Thus, score interpretation is individual to the test taker and/or practitioner and test taker results cannot be compared to those of other test takers. In the case of self-help and self-guided learning e.g. using printed publications and online tests, the interpretation of the results relies largely on the meaning of the activity and results for the individual and on their capacity for self-integration. Where such tests are undertaken by individuals or groups in a structured environment e.g., learning activities accompanied by a teacher/trainer, or counsellor, or psychologist, the reliability and validity of the response

interpretations depend on the practitioners' observational and interpretive skills in addition to their client communications skills, and also on the self-integration capacity of the individual.

2.2 Towards standards for the development of non-psychometric tests

While industry standards or national or international guidelines do not exist for non-psychometric tests, some test developers have drawn on some of the *Standards* for psychometric tests to guide their test construction. A good example of this is provided by McMahon, Watson, and Patton (2002). In developing an assessment tool, My System of Career Influences (MSCI), they drew upon a theory of career development, the Systems Theory Framework (STF; McMahon and Patton, 1999) to establish the tool's construct validity. Content validity was established by the fact that the three authors were career development experts and that two of them had developed the theory. Face validity was established through trialing the tool with a range of career development practitioners. Test items were developed and tested on two different small samples of university students (Australia, South Africa) using a pilot version of the test and revised according to the results of the pilot testing. Test instructions were written in such a way that the tool could be a self-guided instrument. Feedback on the tool was sought from the two samples as to age appropriateness, language and terminology, concepts, the diagram, completion time, the structure and layout, the written instructions, the debriefing process, and the tool's potential usefulness. An adolescent form of the MSCI was developed on the basis of this feedback and subsequently tested with small samples of university students (South Africa, Australia) and adolescents (socio-economically disadvantaged, some with assessed learning disabilities). Feedback on the instrument was sought. Some differences emerged between the feedback of the adults and adolescents on ease of language and of instructions, and on relating to the concepts of career and of systemic thinking. The authors concluded that the process of development and refinement through trials and feedback now allow it to be used on a larger scale with adolescents.

McMahon, Patton, and Watson (2003) put forward the following suggestions for the development of 'qualitative' career assessment tools:

- Ground the assessment tool in theory
- Pilot test the development of the tool e.g. item development and selection; use of adequate samples of ultimate target groups to test the items; obtain feedback from participants and practitioners in the pilot test on the tool's relevance and usefulness, and on the administration process including the optimal time required for completion of the assessment
- Promote a holistic approach to client's needs, concerns, issues rather than a narrow focus on the object of the tool
- Provide the client with personalized written instructions that encourage the client's engagement with the tool
- Provide readable and easily understood instructions
- Structure the tool in a logical flow and sequence of easy steps for the client to take

- Use a focused and flexible process whereby, for example, clients are allowed to choose whether to complete the exercise or parts of the exercise if they do not find it personally meaningful
- Encourage the involvement of the practitioner and the client as a cooperative process
- Hold a debriefing session to allow client feedback on their experience of both the tool and the process.

In her 2019 review, McMahon further addressed the issue of guidelines for the development of non-psychometric/qualitative tests by proposing the application of principles of qualitative research in general to the development and use of qualitative career learning tools, in particular, the principle of trustworthiness and its criteria of credibility, dependability, confirmability, and transferability (Lincoln and Guba, 1982). Credibility refers to the truth and accuracy of results – the data is representative of participants and their experiences; dependability concerns the consistency in results when the methods are applied by other researchers/practitioners to similar populations and contexts; transferability refers to the applicability of the methods and results to similar contexts, populations, and circumstances; and confirmability concerns how accurately the results reflect the participants responses as opposed to the researcher's/practitioner's bias. Confirmability is essentially established through the first three criteria.

2.3 The value of psychometric and non-psychometric test results

It is important to note that information obtained through career assessment tools, both psychometric and non-psychometric, is not the primary nor the sole basis for career decisions (Osborn and Zunker, 2016). Self-knowledge or learning about oneself is a step in a career learning process. Being able to self-integrate this information can also be challenging. Additionally, there are many external influences that impact on one's decision-making e.g. socioeconomic status, family finances, geographical location, cultural and family influences, locally or regionally available learning and work opportunities, knowledge of the world of work and occupations, skills and qualifications acquired, work experience or the lack of it. McMahon and Patton (2002) propose a set of guidelines for practitioners using non-psychometric tests in career counselling in a practitioner-led intervention.

2.4 Summary

Non-psychometric tests refer to questionnaires, checklists, card-sorts, drawings etc. that are used by career practitioners, other professionals, and by the public, to help people develop self-knowledge that may be useful for different types of decision-making e.g. personal, learning, work related, relationships etc. Many of them are stand-alone, self-help instruments that can be found in written/printed form, computerized, or online. They are self-administered and self-interpreted. Non-psychometric tests in general are not developed following any specific national standards as per psychometric tests. Indeed, as far as this author could ascertain, no such standards exist in any country worldwide. The topic of non-psychometric tests in career learning has attracted very little academic interest, with the exception of Prof. Mary McMahon (Australia) and her international research colleagues. Part of the problem lies in the lack of any clear professional definition of what constitutes a non-psychometric test and in the amorphous nature

of terminology used to describe such tests. McMahon and colleagues (2003) have made some suggestions of guidelines for non-psychometric test development, but they are very much a starting point.

Among the advantages of non-psychometric tests are their low development costs, their ease of accessibility by the public, their flexibility and adaptability for use with diverse population groups e.g. life experience, age, ethnicity, culture, learning disability and other types of disability, and learning style, and their stand alone nature for test administration and interpretation. The downside of non-psychometric tests includes their lack of scientific rigour – lack of statistical estimates of reliability and validity, norms, expectancy tables, interpretative charts etc. that give meaning to test scores. When used as a stand-alone self-help instrument, test score interpretation is in the hands of the test-taker. When such test taking is supervised by a practitioner, test score interpretation depends on the knowledge and skills of the practitioner. Test-takers differ in their capacity to integrate new information about themselves. In general, in a stand-alone situation, the people who benefit most from self-interpretations of test scores are those who already have good self-knowledge.

CHAPTER 3 ONLINE TESTING CONSIDERATIONS

Traditionally, psychometric and non-psychometric testing have used a written (paper and pencil) or pictorial or card format with individual hand scoring by a practitioner, moving in the late 1960s to a computerized format with batch machine scoring possibilities and later to personal computer usage (individual and networked) and individual computerized scoring. Such test-taking involved some form of professional accompaniment to assist test administration, scoring, and interpretation of the results though the shift to personal computer encouraged and sometimes promoted self-directed search, particularly in the field of career learning. There, computerized tests could be taken as stand-alone instruments or as part of a computerized career system that included databases of occupations and of learning pathways to these. Algorithms were developed to link test score results to lists of occupations and learning pathways, and to generate a report or profile for the practitioner and/or the test taker. This represented a huge advance in terms of reducing the time spent by career practitioners on test administration and information giving and increasing their time to provide more specialized help or to assist more clients. Herr et al. (2004) reviewed the research on computerized testing for career learning and concluded that algorithm-derived computer test interpretations functioned more efficiently, accurately, and reliably than average and below average clinician interpretations. They emphasised the need for practitioners, from both professional and ethical standpoints, to have an in-depth knowledge of computerized tests that they use, their underlying constructs, their psychometric properties, their appropriate population applications, score interpretations, and their value as just one source of career learning information to assist client decision-making. They also noted that competence in test score interpretation, as part of a practitioner's work, was the most under-researched area of career counselling while paradoxically being one of the most continually debated aspects of a practitioner's role.

3.1 Online testing

Since the 1990s, with the development of the worldwide web and advances in ICT, there has been a significant move to an online testing format, in some ways mirroring advances in computerized testing, both for stand-alone and as part of integrated career planning systems, but with many additional features such as virtual online support. Osborn and Zunker (2016) observe that the developers of such career systems placed less emphasis on the psychometric properties (e.g. validity, reliability) of the tests and more on their function as a learning tool, including providing a starting point for career learning and exploration.

The provision and use of online psychometric and non-psychometric tests are attractive for many reasons:

- The provision of online tests has the potential to reach a very large and wide national audience
- A test taker can log on to a website that has a test from any geographic location that has internet connectivity in order to do the test, and at a time of his/her convenience 24/7, using a computer or smart phone or tablet
- Online tests promote self-knowledge, important for self-concept development and for relationships, providing immediate learning feedback which can increase motivation for further learning and exploration activity
- The accurate compilation of test scores is assured
- Feedback can be provided on responses to individual items, to groups of items, and to the test as a whole
- Feedback can be provided in terms of raw scores, or grade or percentile ranks where norms to enable individual or group comparisons have been established
- Test developers can change question items easily if they are found not to be appropriate
- Advances in the use of Gameful Design make for easier and customized online learning assessment of any kind.

3.2 The challenges of online test usage

The use of online testing also has some significant downsides:

- It is difficult for a test taker and practitioner to judge the quality and value of a test, of test score interpretation, and of test score reports and profiles, unless the test is already a well-known and well-researched test that has been made available online.
- Often tests are provided online with good marketing information but without any information about the developers, the test development, and test evaluation data, thus making it difficult for public and professionals to make an informed selection of a test. It's a case of 'buyer beware'.

- Consumer protection is weak. While practitioners who have appropriate psychometric training and who are members of relevant professional associations have sufficient knowledge and experience to make some judgement about online tests and their source, no internationally or nationally recognized quality mark exists for both psychometric and non-psychometric tests that would assist the public in making informed choices.

-The proliferation of online tests is a professional challenge. Professional bodies and associations do not have the capacity to patrol the internet, nor the power to attach warnings to online tests deemed of dubious quality, nor can they control the choices and behaviour of the public in the private selection and use of online tests.

-Equity of access⁴ to online tests (also an issue for printed and computerized tests): access to online tests by the public requires (i) financial capacity to purchase, borrow, and use the relevant hardware - computer, tablet, or smart phone, and software (ii) broadband quality to enable such access, and (iii) digital competence (intellectual and physical) to use hardware and software. The non-fulfillment of any or all of these three conditions may deprive certain and significant groups in society of access to and use of online tests.

-Individuals differ in their capacities to source online tests, and to interpret, evaluate and apply information derived from test scores. Gustad and Tuma (1957) found that people who benefit most from the self-knowledge gained from tests are those who have more accurate self-knowledge before taking a test. Thus, the benefits of self-directed learning from online tests are not spread equally across the population, with some people requiring differing levels of practitioner support to enable them to self-integrate the knowledge gained from the test results.

-The use of online tests cannot guarantee standard conditions of test taking. Variances in test scores may result from variances in test taking conditions.

-The time required to complete an online questionnaire is a factor in test completion and answer quality (Galesic & Bosnjak, 2009); the shorter, the better.

-The ease of website navigation and the attractiveness of features of the presentation of online tests e.g. use of emoji anchors vs traditional scale anchors, may also affect online test participation.

3.2 Standards and online testing

The *Standards* (AERA, 2014) draw attention to additional specific responsibilities of test developers and administrators as concern computerized and online test usage. These include:

-the use of appropriate hardware and software, and taking into account the specific needs and test taking capabilities of different groups of test takers

-instruction and practice exercises, as appropriate, to help the test taker to use the hardware and software

⁴ See ELGPN (2015) *Guideline 9 – ICT in lifelong guidance*. In Guidelines for Policies and Systems for Lifelong Guidance – a Reference Framework for the EU and for the Commission. Jyväskylä: ELGPN

-in practitioner-supported counselling, clinical, and educational settings, practitioners using computerized or online tests should draw on other relevant test taker information, as appropriate, to support the interpretation of test scores

-assurance of continuance of internet-connectivity during test taking

-security, especially from hacking, of both the test content and of test score results

-a description of the general characteristics of scoring algorithms when automated algorithms are used to generate a report or profile.

In response to the increased use of online testing, the International Test Commission (ITC, 2006) has drawn up a set of guidelines to assist the development, distribution, use, administration and taking of computer based (CBT) and internet tests. These guidelines address:

-technological issues

-quality issues

-levels of control

-security and safeguarding privacy.

The guidelines for each of these areas are subdivided and grouped differently for test developers, test publishers, and test users. These are viewed as the main stakeholders, but others are also referenced: the trainers of test users; the test takers and their families; relevant professional bodies and associations; and policymakers and legislators.

Their content is based on and reflects existing guidelines, codes of practice, standards, research papers, and international consultation. The Guidelines are intended to provide a basis for comparison for existing national guidelines, benchmarking, support their development, and consistency across national boundaries. As in the case of the *Standards*, the pertinence and application of different standards depends on whether decisions based on test scores are high stakes or low stakes.

The ITC guidelines make a useful distinction between different modes of test administration (open, controlled, supervised, and managed). They also highlight contextual factors that need to be taken into account in the national application of such guidelines. These include social, cultural, linguistic, institutional, political, legal, existing national guidelines, test settings, individual vs group tests, the primary recipients of the test results, use of test results (high, low stakes), and whether the accuracy of score interpretations can be checked and/or amended in the light of other available information.

There are some ITC guidelines that are additional to the *Standards*. These include:

-Where possible, encourage test users to collect biographical data on test-takers in order to monitor the number of people from protected/minority groups who take any CBT/Internet test.

-Where unequal access to CBT/Internet tests may occur, recommend that test users make alternative forms of assessment available.

-Assess differential items functioning (DIF) and inform test users of any evidence regarding DIF for different test-taker groups.

3.3 National standards for CBT and online career assessments

Some national associations for career practitioners have developed guidelines for CBT tests, many of which are also applicable to Internet tests. The National Association for Career Development (NCDCA) in the USA uses five categories of software evaluation criteria to help test-users review career software programs. These are:

-information: relevance to the audience, appropriate language, organization of the information, and information quality

-career development process: the compatibility of the program with important career development principles

-interaction with the program: test-taker's interaction with the program, the objectives and features of the program, and the user's analysis of it.

- Technical aspects of the hardware, software and materials

-Support: support for professionals who implement the programme: written materials, staff training, service, and cost.

Some of the 67 criteria apply to any career development software program: nondiscriminatory language, current and valid information, test-taker control of decision making, programme reliability, and availability of technical assistance. Other criteria (e.g., inclusion of test results or appropriateness for small group use) apply to features which may not be important for a particular type of programme.

3.4 Professional association initiative (national)

In a novel approach, the NCDCA has drawn up a list of internet sites for career planning, that are "*free, current and credible*", based on information received from its members. The criteria for website inclusion on the list are as follows:

-Follow the mission of NCDCA which is "to inspire and empower individuals to achieve their career and life goals" and may not violate the NCDCA non-discrimination policy.

-Be viewable by anyone with Internet access without any major or substantial services requiring payment (in other words the critical information is free, and the site does not try to entice users into paying for fee-based services)

-Contain accurate, factual, unbiased, and current information

-Have strong and obvious relevance to career development

-Be easy and clear to navigate

-Provide information of use to the client

-Be professionally prepared and executed (embedded links work, presentation is professional, grammar correct, appropriate use of images, no typos, extra resources and assistance provided when appropriate).

-Display content that is appropriate and fair to all cultures and genders

-Use professional email addresses for contact purposes.

The following eight features of such websites are examined:

- Self-assessments
- General occupational information
- Industry and occupation-specific information
- Education topics
- Employment trends
- Job search
- Special populations
- Videos.

The self-assessments tests include personality, interests, work values, skills, and career planning. The criteria for website inclusion of sites offering these tests make no reference to the presence or absence of psychometric qualities. However, they do mention strong and obvious relevance (validity), appropriate content (validity), non-discrimination, and culture and gender fairness.

This NCDA initiative offers an informal informed and bottom-up approach to engage career practitioners in identifying the value of different career planning websites, including self-assessment tools, and in disseminating their findings.

The NCDA has recently published the 7th edition of its publication *A Comprehensive Guide to Career Assessment* (Stoltz and Barclay, Eds., 2019). The book provides foundational knowledge of career assessment and 27 peer reviews of assessments, which are mainly of USA origin. The reviews are separated into four categories of career assessments: quantitative (14), qualitative (3), research (7), and open source (3).

3.5 Summary

The internet has provided a new setting for test-taking and indeed for test proliferation. This presents many opportunities and challenges for test-takers, professionals, test publishers and distributors, test developers, professional bodies, institutions, and policy and systems developers. In one way, the internet allows tests to more accessible to the general public. In another way, it puts the general public more at risk as there is no way for them to know what a good test is. Additionally, some people have more capacity than others to self-integrate their interpretations of test scores derived from self-assessments.

Attempts have been made at national and international levels to develop guidelines to address the complexity of issues that online testing presents. These include addressing hardware, software, connectivity, and security issues. A key question for all guidelines and standards is whether the interpretation of test scores is high stakes or low stakes as these also have implications for the control, supervision, and management of test-taking. Contextual factors e.g. social, cultural, linguistic, are important in the national application of international guidelines.

Some national career practitioners' associations, such as the NCDA in the USA, have developed guidelines to enable test users to evaluate career software programs that include tests. The NCDA has drawn up a list of 'free, current, and credible' websites based on members'

recommendations and according to a set of criteria that include gender and culture fairness. Self-assessments (e.g. personality, interests, work values, skills, career planning) are examined as one website feature and there the general evaluation criteria for websites apply: relevance, appropriate content, and gender and culture fairness. There is no reference to the psychometric qualities of these tests. The NCDA also publishes a comprehensive guide to career assessments for its members which does address such issues. There is no minimum educational requirement for NCDA regular membership. In order to become a professional member, one must hold a master's degree in counselling or a related field.

The NCDA model of test evaluation, either as part of CBT and online career planning systems or as stand-alone tests, offers a different approach to educating test-users on the value of different self-assessment tests. It is less institutional bound, more collegial, more flexible, and perhaps more appropriate where the interpretation of test scores is low stakes.

CHAPTER 4 THE KNOWLEDGE AND COMPETENCE OF CAREER PRACTITIONERS IN ASSESSMENT

A common theme that runs across all test usage standards and codes of ethics for users of psychometric and non-psychometric tests is the knowledge and competence of practitioners (users) for using tests and interpreting test scores as part of their work-role. This applies in particular to controlled, supervised, and managed settings of test administration but also applies where online and offline assistance is offered to people who have undertaken non-psychometric self-assessments. It is clear that the higher the stakes of decision-making based on test scores, the greater the knowledge and score interpretation skills required.

Countries vary hugely in terms of professional qualification requirements for persons wishing to exercise a role as a career practitioner with more regulation existing in the education sector than in the labour market sector (McCarthy, 2002; CEDEFOP, 2009). There are also significant differences in training method and content within countries depending on the higher education institution that provides such training, and which university department (education, psychology, sociology) within the institution is primarily responsible for training. Countries try to reduce the effects of such variances in different ways. In Ireland, for example, the Department of Education and Skills (DES, 2016) has set down a framework of knowledge, competence, and skills that career practitioners who wish to work in the education sector have to learn and which universities have to teach trainees. In the one year full-time (or 2 year part-time equivalent) postgraduate diploma/masters, the following areas of competence in assessment are set down by the Irish Department of Education and Skills:

Area of competence	Demonstrate knowledge and	Demonstrate ability to	Demonstrate ability to	Demonstrate ability to	Demonstrate ability to	Demonstrate

	understandi ng of:					
Psycho- metric Testing	-Theories, applications and limitations of psychometric testing -Statistical methods for psychometric testing -Legislation relating to data protection considerations when employing psychometric testing in educational settings	-Critically evaluate the validity, suitability and reliability of ability, attainment and interest tests (including online tests) taking account of the context e.g. information needs, cultural differences, educational disadvantage, literacy and numeracy skills, and special educational needs -Select and apply tests appropriate to requirements and one's own professional competence • Employ appropriate	-Select and administer tests appropriately in a range of contexts and to individuals and groups with diverse needs -Provide feedback appropriately and in a meaningful way to test takers	-Employ testing in accordance with one's own role and competence	-Critically evaluate one's own practice and competence in relation to test administration, scoring, interpretation and feedback - Identify own learning needs and take steps to address them	- Understand ing the scope of practice and boundaries of one's own professional competence

record
keeping and
confidentiality • Inform
test takers of
the
purposes,
and
limitations
of tests and
the validity
of the uses
to be made
of the
findings

• Inform
test takers of
the data
protection
issues
relating to
online tests •
Obtain
informed
consent
(confidential
ity and data
protection)
prior to
administrati
on • Analyse
and interpret
test results

The university training institutions are required to deliver thirty hours of psychometric training to trainees in order for them to obtain certification in psychometric training from the Psychological Society of Ireland.

An alternative approach is used in the UK where the professional career practitioner association, the Career Development Institute (CDI), is responsible for the National Occupational Standard (NOS) for career practitioners and the awarding body for the Qualification in Career Development (QCD). In these roles, the CDI (2016, 2021) sets down standards for the

competence of career practitioners to assist the universities that deliver the programme to meet the requirements of QCD. While both the NOS and the QCD use terms such as ‘techniques’, ‘interventions’, and ‘tools’ to assist clients in career exploration, the terms ‘psychometric’ and ‘tests’ are noticeably absent from both documents, reflecting a move away from attaching importance to tests as tools in career learning, and a move from a psychological to an educational paradigm in career learning itself. Training in psychometrics does not form part of the initial training of career practitioners in the UK. It is a post-training professional development, paid for by the employers or by the practitioners themselves, leading to British Psychological Society certification.

The examples of Ireland and the UK above illustrate differences in the importance attached to psychometric and non-psychometric testing in career learning and in the training of career practitioners by two neighbouring Anglophone countries. While the history of publicly funded career learning provision of both countries dates back to the 1960s, it is clear that cultural, social, philosophical, linguistic, legal, institutional, and political factors have had a strong influence on the content of the training of career practitioners and consequently on the career learning activities, including self-assessments, that are offered to the public.

From an EU overview perspective, CEDEFOP (2009) undertook a survey of the training and qualifications of career practitioners in 21 Member States. They noted that in some countries, career practitioner was a specialist occupation undertaken by persons with specific training while in other countries career guidance activities were a subspecialism within another role e.g. teacher, public employment service official, psychologist, or carried out by those who have obtained career guidance posts on the basis of general qualifications or of extensive life experience that is seen as relevant to career guidance. The authors concluded that the equivalent of one year of full-time higher education was an appropriate benchmark for national discussions in EU countries on the desirable extent of specialised training.

The CEDEFOP authors devised a competence framework for career practitioner, comprising of six foundation competences, six client-interaction competences, and seven supporting competences. Client-interaction competences included undertaking career development activities such as enabling clients’ self-understanding through assessment of skills, aptitudes, interests, life history, achievements and setbacks, hopes and fears, health and personal circumstances. The competence “Conduct and enable assessment” includes (i) the usage by practitioners of checklists, inventories, lifelines; (ii) the competence to administer formal assessments e.g. psychological tests such as aptitude and ability; (iii) knowledge about norms, profiles and reports, reliability and validity of assessment tools; and (iv) understanding the implications and relevance of assessment results for clients’ plans and strategies. Fostering client self-understanding is viewed as a key outcome of the work of a career practitioner and is achieved with the assistance of informal and formal assessment methods. The survey report makes clear that assessment as main task of a career practitioner is moderated by the context and conditions of a career practitioner’s work-role and context.

While one would rightly expect that the content of training of career practitioners, at a minimum, would be work-role and work context bound, and based on job analysis and job reviews of career

practitioners, it is often the case that the training method and content is decided mainly or solely by the trainers (McCarthy, 2002), leading to in-country variances in the outcomes of training and of the subsequent job performance of career practitioners.

Summary The training of career practitioners in the use of formal and informal assessments for career learning is very much related to (i) institutional factors such as work-role and work-context, (ii) the philosophies that underpin training (psychological vs educational vs sociological paradigm), and (iii) how decisions about the content of training are made and by whom. While these factors contribute to variances in the competences and practices of career practitioners, there appears to be general agreement that enabling client self-understanding is a key outcome of a career practitioner's work, and that formal and informal assessments have a key role to play to make such an outcome happen. Thus, career practitioners need the requisite knowledge and competence to assist the public to use online tests, to select and administer such tests as part of blended career learning for clients, to interpret test scores as appropriate, and to communicate these interpretations to clients both online and offline.

CHAPTER 5 ETHICS OF TESTING FOR CAREER PRACTITIONERS

When tests (psychometric and non-psychometric) are administered by test users (e.g. career practitioners) in controlled, supervised and managed settings, professional bodies e.g. the American Counselling Association (ACA, 2014), have developed ethical guidelines to support practitioners in such assessment activities. The ACA guidelines refer to:

- Assessment purpose and duty of care to test-takers
- Competence to use and interpret test instruments
- Informed consent
- Release of data
- Diagnosis
- Instrument selection
- Conditions of assessment administration
- Multi-cultural issues/diversity in assessment
- Scoring and interpretation of assessments
- Security of instruments and scores
- Obsolete assessments and out-of-date results
- Assessment construction
- Forensic evaluation considerations for legal proceedings.

These ACA guidelines form the basis of the NCDA Code of Ethics (2015) section on evaluation, assessment, and interpretation.

The NCDA code also includes a section on the provision of online services that includes:

- Knowledge and skills regarding related technical, ethical, and legal issues
- Informed consent and security
- Client verification
- Providing career services online (including assessment)
- Records
- Web maintenance and technology development
- Social media.

Some additional guidelines on assessments are provided for online assessments:

- for psychometric tests: determine that the psychometric properties of online assessments are the equivalent of those in printed form
- determine if the assessments have been validated for self-help use
- assure that appropriate intervention is provided before and after completion of the assessment resource if the resource has not been validated for self-help use
- refer clients to qualified career professionals in his or her geographic area, if there is evidence that the client does not understand the assessment results.

Codes of ethics are intended to support the work of members of professional associations (both regular and professional members in the case of NCDA), to enable them to make professional decisions about all aspects of their work with clients and with other professionals, that can bear public scrutiny. They also act as a guide to consumers as to what they can expect as professional behaviour from members of the association.

Summary: Career practitioners have a duty of care to their clients and this includes (i) assisting clients to make choices of psychometric and/or non-psychometric tools to facilitate their self-understanding in the context of career planning and decision-making, (ii) administering, scoring, and interpreting test results, and (iii) ensuring the security of the tools and of the test scores. Professional bodies have set down good practice in career assessment through codes of ethics which members are requested to adhere to. Such codes of ethics provide a basic protection for career practitioners and their clients in assisted, controlled, supervised, and managed assessment settings.

CHAPTER 6 REVIEW OF FREE ONLINE PUBLICLY FUNDED CAREER ASSESSMENT TESTS ON GOVERNMENT SUPPORTED WEBSITES IN EUROPEAN COUNTRIES AND NATIONS⁵

Since the 1990s governments in the European Union (EU) have provided funding for online career services, mainly through the ministries of education and labour. Such services have included online stand-alone self-assessment tests. As part of the present task, this author has undertaken a survey of what currently exists in EU countries, based on an inventory of guidance systems and practices (CEDEFOP, 2020). The results are provided in Annex 1. The author sought to discover for each EU country:

- the responsible government ministry or agency plus the website web link
- the subject of the test or questionnaire
- the functions of the test or questionnaire
- the psychometric nature (or not) of the test
- the support given to the test taker (online/offline, synchronous/asynchronous).

Responses from twenty-four countries and nations represented in CEDEFOP's CareersNet were examined and checked with national correspondents and national experts as appropriate. Additional information was sought from Croatia, giving a total of 25 possible country responses.

6.1 Free vs paid for tests

The results (Annex 2) show that 22 of the 25 countries and nations provide free publicly funded online career self-assessment tools (three countries do not provide online tests). In addition, in two of these 22 countries, France (partner organisations of the public employment service) and Netherlands (partner organisations in the education and employment sectors), a first level of service by partner organisations including online tests are free but the second level of service including more extensive psychometric tests have to be paid for by the public.

6.2 Terminology

In general, countries whose tools were of a psychometric nature tended to call them 'tests' while those of a non-psychometric nature were also called 'questionnaires' or 'tools' or 'game' or 'quiz'. On some websites, some of those five terms were used interchangeably, reflecting cultural meaning and appropriateness of the terms for engaging the intended target group.

6.3 Test focus and function

The most common subject of such tests were occupational interests/preferences, but some countries, in addition, provide self-assessment tests of personality, work values, abilities, skills, competences, and study and career planning, mostly as separate instruments. The main functions of the tests are to support career learning: self-appraisal and career exploration. However, the

⁵ Please refer to Methodological Note in Annex Two

career planning, decision-making, and study skills tools could be viewed more as screening or problem identification instruments, to be used in a context where career practitioners are available to help clients to progress towards readiness to decide and plan.

6.4 Test result job matching

In many instances, the test results (especially of interests) are linked by an algorithm to an occupations database, thus enabling test-takers to view occupations that reflect their test scores or to compare their test scores to any occupation in the database (e.g. Finland www.asiointi.mol.fi). In some countries, interest test results are linked by an algorithm to a database of education and training programmes, thus enabling test-takers to view learning programmes that reflect their test scores (e.g. Ireland www.qualifax.ie). In Finland (www.asiointi.mol.fi), interest test results are linked by algorithms to both occupations and learning opportunities databases. Where different psychometric test results are combined to give a matching occupations list (e.g. Germany), more weighting is given to the skills results than to the interests and preferences results. Occupational taxonomies and derived skills descriptions enable the matching of skills assessments results with occupations (e.g. Germany, France, Scotland).

6.5 Psychometric? Theory-based?

In a majority of countries, the tests are non-psychometric in nature while in some countries (e.g. Austria, France), one can find both psychometric and non-psychometric tests. Holland's career choice theory (or varieties of his classification system), matching one's personality type with similar categorization of occupations, are the basis of several of those instruments, both psychometric (e.g. France, Germany, Greece) and non-psychometric (e.g. Slovenia, Sweden). Given the non-existence of guidelines for the development of non-psychometric tests, it was interesting to discover that in the education sector in France, one interest test was developed by a group of career practitioners based on their own knowledge, experience, and common sense, while another was developed for a partner organization in the labour market sector by someone who also used their 'common sense' and who was not aware of the existence of the science of psychometrics!

6.6 Standards?

Psychometric test developers in France (and other European countries) use the *Standards* (AERA et al., 2014) as a benchmark. The five top test publishers in France have formed an association of test publishers (AET) and agreed and published a *Charter for Personal Evaluation Tools* (2009) which includes three sets of principles concerning (i) the construction and validation of a test; (ii) the involvement of the national and international scientific community (university and other researchers) in their development; and (iii) the usage of the tests only by trained professionals. The objectives of the charter include to promote the scientific nature of the psychometric test publishers' work and to defend their commercial interests. No such association or charter exists for developers and publishers of non-psychometric tests.

6.7 Test-taker target group differentiation

Some countries (e.g. France) use a targeted approach e.g. different tests for different categories of youth and for adults and according to administrative responsibility e.g. education (students) or employment (jobseekers) ministry, while others (e.g. Germany, Greece) have a uniform approach – the same test for youth and adults. However, in Germany, biographic information given at the time of website registration dictates the content and level of difficulty of the question items presented to test-takers, distinguishing between vocational training bound and university bound students. Of particular interest are the Diagoriente tools in France (<https://app.diagorientebeta.gouv.fr/>) which have been developed by a multidisciplinary team for use with NEETs (youth not in education, employment, and training) based on an online pedagogical approach. An adult version of the tool is due to be published in 2021.

6.8 Smart phone test version

Pole Emploi (Public Employment Service) in France has also developed Quiz Pro, a smartphone version of its interest test. It was originally developed as a psychometric paper and pencil test, used only by work psychologists in a client accompanied setting. Subsequently, the test was transformed to an online version and other staff have been trained internally to use the test in their daily work with the unemployed. Likewise, a smart phone version of interests test has been developed in Finland.

6.9 Access to tests

Some countries require test-takers to open an account (or use their unemployment registration number or national ID number e.g. Estonia) in order to take the test or in order for test takers to keep a record of their results online or to enable them to take a pause during test taking while others do not have this requirement. In Germany, account data is used for developing norms for tests but test-takers are allowed to open anonymous accounts. There is restricted access to tests in some countries e.g. Slovenia, where pupils and students can only obtain access to psychometric tests through the school counsellor or through the national careers centre (Public Employment Service); likewise only registered unemployed adults can obtain access to psychometric tests. Likewise, in Estonia, ability and personality tests are only accessible through a career counselling interview (face to face or remote) at the public employment service.

6.10 Test usage and evaluation

Data such as logins, logouts, test completion rates, provide basic data on test usage that can be used for preliminary evaluation. An account registration requirement for access to tests can provide additional information about test-taker profiles but it can be difficult to balance attracting different groups of the public to use the tests against registration requirements and privacy concerns about the usage of test-taker profile data. Views of test takers on the perceived usefulness of the tests are gathered in different ways e.g. Scotland, as part of an annual careers services survey, or through an online follow-up email survey where test-takers have accounts, through dedicated customer surveys, through practitioner feedback, and through observation of test-takers behaviour during accompanied career guidance individual and group sessions. In

Germany, continuous user evaluation feedback is sought through monitoring test usage, collecting and analysing test-takers ratings of the test and test-takers comments (feedback option with the test), through feedback from career practitioners and teachers, and from observing the behaviour of test-takers in real time.

6.11 Data retention

Data retention policies and terms and conditions of website usage vary within and across countries and agencies. Addressing test-taker concerns about the security and onward usage of test takers' data, one Finnish test website (www.crear.fi) states: 'We do not keep your answers nor results. You can use this tool without any pressure'.

6.12 Different language versions

Some countries, e.g. Finland, offer two language versions of the same test (Finnish, Swedish), and more recently four versions of the same test (Finnish, Swedish, plain Finnish, and audio www.crear.fi). Estonia offers the interest test in Estonian, Russian, and English (<https://www.tripod.ee/>).

6.13 Test-taker support

In most countries, free online career assessments are offered as stand-alone instruments without synchronous or asynchronous practitioner accompaniment. The general provision of free publicly funded online tests is based on several assumptions, including that all online test-takers are autonomous learners. But as pointed out in Section 3.1 above, the benefits of self-directed learning from online tests are not spread equally across the population, with different groups of the public requiring differing levels of practitioner support to enable them to self-integrate the knowledge gained from the test result. In Germany, the online tests are accompanied by a qualifying statement (<https://www.arbeitsagentur.de/bildung/selbsterkundungstool-faq-allgemein>) that while the tools provide a basis for initial career exploration, they are not a substitute for careers advice given by a careers practitioner. Career practitioners visit schools to inform teachers and students of the online and career counselling resources available to them and teachers use the tools during occupational orientation classes. The statement provides a weblink to a request for an interview with public employment service personnel. Where accompaniment is provided, countries vary in the type of accompanying support given to test-takers. For example, Pole Emploi (France) organizes different categories of individual and group support for users: light touch support for the more autonomous learners; guided support for those who need some assistance e.g. once off help; and enhanced support for those with the biggest labour market entry challenges. To better support the first category, more online tools were developed and/or presented by Pole Emploi partner organisations (Emploi Store). Practitioners supporting the other two categories of users have recognized the value of such tools and have incorporated them in their practice. Accompanying support is both synchronous (e.g. online chat or telephone or during a real-time individual or group intervention) and asynchronous (contacting the practitioner by email subsequent to taking a test).

6.14 Conclusion

Free publicly funded online career learning assessment tests are provided in most European countries, mainly as stand-alone self-directed instruments. The majority of such tests are non-psychometric in nature. No guidelines or standards exist for the development of such tests, but theoretical and career expert, and more recently, multidisciplinary input, are evident in their development. More than half of the countries have separate tests for youth and for adults. Many tests allow test-taker results to be matched with education and training pathways and with occupations. Funding for youth tests comes mainly but not exclusively from the ministry of education while funding for tests for adults comes mainly but not exclusively from the ministry of labour. While online tests are standalone tools for self-appraisal and career exploration, countries offer varying degrees of support to test-takers. Online registration allows for usage and test-taker evaluation. There is an increasing recognition that online test design and development requires the input of multidisciplinary teams and the use of Gameful Design principles.

CHAPTER 7 TOWARDS A SET OF GUIDELINES FOR THE DEVELOPMENT OF NON-PSYCHOMETRIC CAREER LEARNING TESTS

7.1 Introduction

The science of psychometrics concerns the precision of mental measurements and the meaningfulness of those measurement scores. Such precision and meaningfulness are really important where measurement scores are used for decision-making and for inter-individual comparisons such as in workplace recruitment, promotion, prediction (workplace, academic), and in clinical diagnosis and treatment. These decision-making activities have high stakes consequences. Psychometric measurements are also helpful for self-understanding and self-appraisal. The use of psychometrics in ipsative assessments that yield intraindividual difference scores can clarify for an individual the relative strength of a particular construct in comparison to other constructs for that person.

Self-understanding and self-appraisal are lifelong iterative processes which develop, inter-alia, through interaction with many different aspects of one's environment. Non-psychometric assessments such as comments from one's family or friends or work colleagues contribute to developing a sense of self, identity, self-appraisal, and self-understanding. They are often competing voices from different significant sources. Likewise, non-psychometric tests, questionnaires, checklists, games, roleplays, drawings etc. can contribute to self-understanding and self-appraisal of one's personal qualities, interests, skills, values, career planning ability, study habits etc. They provide additional sources of self-information in a context of many different possible informants.

Internationally recognized standards and guidelines exist for psychometric measurements in education and psychology (Chapter 1 above). This makes a lot of sense given the high stakes nature of the consequences of decisions made on the basis of psychometric test scores, the commercial business that these tests generate, and the need for customer (e.g. employer,

healthcare system) and consumer (e.g. individual test-taker) and professional (e.g. clinician) protection. The field of non-psychometric assessments does not possess such formal guidelines although attempts have been made to define some of these in research literature (Chapter 2 above). The need for such guidelines is important in the context of publicly funded free online career assessment tests although it is clear, from the review of such online tests in 22 European countries and nations, that government ministries and agencies have developed many useful online career assessment tests in the absence of guidelines. Based on the literature review (Chapters 1 to 5 above), the field research in European countries (Chapter 6), and exchanges with international correspondents and experts, the following are some suggestions for guidelines for non-psychometric career assessment tests.

7.2 Suggestions⁶ for draft guidelines for non-psychometric career learning assessment test development

7.2.1 Target groups:

These suggestions are intended for:

- government ministries and agencies who fund non-psychometric career assessment test development, or purchase or publish such tests or who promote good practice in non-psychometric test development or who set standards for the training of career practitioners who use such tests
- government and government agency and private sector personnel employed as test developers
- private sector commercial companies who publish non-psychometric career assessment tests
- university and other higher education institute personnel whose work involves non-psychometric test development
- career development practitioners who use such tests
- professional associations of career development practitioners
- the general public of test-takers.

7.2.2 Limits of the guidelines

The draft guidelines are intended to act as a compass for good practice in career learning test development and not for regulation (as per the *AERA Standards*, 2014). The acceptability of a test or testing practice does not rest on the literal satisfaction of every guideline. At the same time, guidelines should not be seen in isolation from one another. Some evidence is needed for propositions that underlie test score interpretations. Contextual factors also need to be taken into account in the application of such guidelines such as social, cultural, linguistic, institutional, political, legal, and resource factors; test settings e.g. online, offline, individual and group, self-

⁶ These suggestions are mainly based on the *Standards* (AERA, 2014), the *Guidelines* (ITC, 2006), the proposals of McMahon, Patton, and Watson (2003) and McMahon (2019), and the findings from the survey of 22 European countries.

directed or accompanied; the primary recipients of the test results; use of test results (e.g. formative self-career learning); and whether the score interpretations can be checked and/or amended in the light of other available information.

7.2.3 Draft Guidelines

1. Validity of test scores

(i) Conceptual framework: use a theoretical background and other evidence, if at all possible, to support the interpretation of test scores for the proposed uses of the test. These help to clarify the construct(s) that the test purports to assess. Such an approach enhances test credibility, dependability, confirmability, and transferability, the principles of qualitative research, as advocated by McMahon (2019) (Chapter 2.2 above). In the absence of a theoretical background, provide a very clear description of the construct(s) or domain being assessed that makes sense to the test-taker.

(ii) The construct(s) or domain should be meaningful and understood in the culture and society of the test-taker.

(iii) The test content (test items and internal structure) should be related to the constructs or domain in an obvious way. This mapping and evaluation can be undertaken by persons with expertise in the field (see example of McMahon, Watson, and Patton, 2002, in Chapter 2.2 above) supported by persons from other disciplines e.g. sociology, education. The mapping and evaluation should be documented.

(iv) The test content (test items and structure) should on face value appear to measure what the test claims to measure. This evaluation can be undertaken by persons with expertise and/or through feedback from a sample of the intended test population

(v) If more than one construct is being assessed, the relationships between the constructs should be made clear.

(vi) Specify the population(s) for which the test is intended and the context in which the test scores are to be used and how.

(vii) Test scores: where test scores (sub-scores, totals, or composite scores, profiles) based on test-taker responses are calculated and provided, their interpretation and the rationale and evidence for interpretation should be provided based on (i), (ii), and (v) above. The limits of score interpretation should also be made clear.

(viii) Evidence supporting the validity of test scores should be collected and reported by test developers and publishers. Practitioner feedback from test usage can provide ongoing evaluative data.

2. Reliability of test scores

(i) Generate as many test items as possible for a given construct or domain using a group of experts.

- (ii) Ensure that the sample of test items reflects the universe of possible test items for a given construct or content domain, using a group of experts.
- (iii) Trial a sample of the test items with heterogeneous sample groups of the intended test target group, observe their item response patterns (differential item functioning), and obtain their feedback on the items. (see McMahon, Watson, and Patton, 2002).
- (iv) Trial varying amounts of test items and test-time duration for optimal participant engagement.
- (v) As appropriate, trial separately different formats of a test e.g. audio, large print, different languages, age groups etc.
- (vi) The trial conditions in (iii) to (v) should reflect the ultimate contexts of test-taking e.g. online or computerized or written or assisted or supervised.
- (vii) Educate practitioners about the test and train them in administration, scoring, score interpretation, or in rating when the test requires practitioner administration (including specific conditions of test taking), practitioner test scoring and score interpretation, or subjective practitioner rating.
- (viii) Evidence supporting the reliability of test scores should be collected and reported by test developers and publishers. Practitioner feedback from test usage can provide ongoing evaluative data.

7.3 Fairness issues that can affect the validity of test score interpretation

- (i) In as far as possible, use the principles of universal design in assessment to maximise test-takers' ability to respond to test-item tasks. The characteristics of individuals of the intended test population should be considered at all stages of test development, administration, scoring, and interpretation. Such characteristics include gender, age, ethnicity, race, socioeconomic status, linguistic and cultural background, test-taking ability (relating to forms of disability), individually or in combination, and any others, that can limit the validity of test score interpretation.
- (ii) Ensure that test item task complexity e.g. vocabulary used, does not disadvantage the targeted test population.
- (iii) Review the test items for content quality, clarity, and construct relevance, and document the review findings.
- (iv) Test developers and publishers should consider existing laws or statutes, if any, that govern fairness in the content and application of non-psychometric tests.
- (v) In so far as possible, developers and publishers should provide practitioners with clear standardized test administration instructions including specified time limits, specified test-taking conditions e.g. room arrangements, and specified practitioner assistance as appropriate.

(vi) Ensure, in so far as possible, that the characteristics of the accompanying practitioner, when such is provided/intended for test administration, scoring, and interpretation, match those of the test-taker, thus reducing the risks to test fairness arising from practitioner characteristics.

(vii) Take into account the degree of a test-taker's bilingualism or multilingualism that may impact on their test responses.

(viii) To optimize test-taker participation, engagement, and performance, provide tests in the test-takers target group's native language.

(ix) Use a multidisciplinary team of independent experts to review the cultural appropriateness of language, illustrations, graphics, photos, and other representations in the test item contents.

(x) When the internet or computerized version of a test is not available due to the absence of or inadequacy of the required hardware, software, and connectivity, a written version is made available to groups of the intended target group.

7.4 Test design and development specifications

(i) Test developers should create test specifications that describe the rationale and intended uses of the test, the choices of test content, format, length, duration, delivery mode(s), administration, scoring, test score interpretation, and reporting.

(ii) Test specifications should be reviewed by relevant external experts including practitioners to evaluate their appropriateness, and the results of the review documented.

(iii) Instructions given to test-takers should be sufficiently clear that test-takers can respond in the appropriate way as intended by the test developer

(iv) Prior to test administration, test-takers should be provided with practice opportunity or sample questions or practice examples that maximises their understanding of how they are expected to respond to test items.

(v) The test developer should provide the practitioner and/or test-taker with procedures for scoring of sufficient detail and clarity that enable the accuracy of scoring.

(vi) For online and computerized testing, when an automated algorithm is used to score test-taker responses, the general characteristics of the scoring algorithm is documented.

(vii) The test developer and publisher should specify the training that practitioners require in order to use (administer, score, interpret) the test.

(viii) When a test undergoes revision, practitioners should be informed of the changes.

7.5 Test scores and administration

(i) Test users (and test-takers as appropriate) should be provided with clear explanations of the meaning, interpretation, and limitations of the test scores.

(ii) Test score interpretations should describe in simple language and as appropriate to the audience of test takers what the test covers, what the scores represent, and how the scores are intended to be used.

(iii) Ensure the security of test score data in line with data retention legal requirements.

(iii) Where appropriate, test users should follow the standard test administration procedures set down by test developers.

(iv) Test developers should provide information on the appropriate environment for test taking to minimise construct-irrelevant variance arising in test scores.

7.6 Test documentation

(i) Test documentation should include the rationale, intended target group, test development procedures, recommended uses, information to support score interpretation, a manual for test users, user qualifications and training, test administration and scoring procedures.

(ii) Scoring instructions and interpretative information be provided in a simple language for test takers where tests are stand-alone, self-directed learning instruments.

7.7 Test users' responsibilities

(i) As appropriate, test users should be trained and qualified to administer the test and interpret the test scores.

(ii) Use professional knowledge in test selection

(iii) Be alert to possible misinterpretations of test scores by test-takers.

(iv) Should not rely solely on internet, electronic device, or computer-generated interpretations of test scores.

(v) Verify the quality of evidence for the validity of computer-generated interpretations of test scores.

(v) Consider the validity of test score interpretation of test takers whose primary language is not the language of the test.

(vi) Ensure the security of test score data of test takers.

(vii) Respect test copyrights.

7.8 Additional considerations for online and computerized test development and usage

(i) Multidisciplinary teams (pedagogy, media, sociology, career behaviour and development) be used for the design and development of online career learning assessment tests.

- (ii) Consider the use of principles of Gameful Design to inform test design and development approaches.
- (iii) Use appropriate hardware and software, and taking into account the specific needs and test taking capabilities of different groups of test takers
- (iv) Provide instruction and practice exercises, as appropriate, to help the test taker to use the hardware and software
- (v) Assure web browser compatibility to access the tests
- (vi) Assure the continuance of internet-connectivity during test taking
- (vii) Assure the security, especially from online hacking, of any recorded or registered test-taker data, of test content and of test score results.
- (viii) Document the general characteristics of scoring algorithms when automated algorithms are used to generate a report or profile.
- (ix) Ensure that documentation is provided on the rationale and basis for interpretation of automatically generated test score reports
- (x) Where possible, collect biographical data on test-takers in order to monitor the number of people from protected/minority groups who take any CBT/Internet test.
- (xi) Where unequal access to CBT/Internet tests may occur, provide alternative forms of assessment e.g. written.

CHAPTER 8 ADDITIONAL RECOMMENDATIONS FOR SOUTH AFRICA

- (i) The South Africa Career Development Association follow the example of NCDA (USA) in drawing up a list of ‘free, current, and credible’ websites based on members’ recommendations that include self-assessments (e.g. personality, interests, work values, skills, career planning) as one website feature and based on general evaluation criteria as used by NCDA: relevance, appropriate content, and gender and culture fairness. Such an initiative would provide a minimum light touch evaluation and advice approach to free online career assessment tests of both a psychometric and non-psychometric nature.
- (ii) DHET pay attention to the new type of test development for NEETs in France (<https://app.diagorientebeta.gouv.fr/>) based on skills identification from social learning and the use of a multidisciplinary approach to test design and development. It is open source.

REFERENCES

- AERA, APA, and NCME (2014) *Standards for educational and psychological testing*. Washington DC: AERA
- American Counseling Association (2014) *Code of Ethics*. Alexandria, VA.: ACA
- Association des Editeurs de Test (2009) *Charte des editeurs d'outils d'evaluation des personnes*. Paris: AET
- Career Development Institute (2020) *Qualification in Career Development – Handbook 2020/2021* Stourbridge, UK: CDI
- Career Development Institute (2014) *Career Development National Occupational Standards*. Stourbridge, UK: CDI
- CEDEFOP (2020) *Inventory of lifelong guidance systems and practices*. <https://www.cedefop.europa.eu/en/publications-and-resources/country-reports/inventory-lifelong-guidance-systems-and-practices>
- CEDEFOP (2009) *Professionalising career guidance: practitioner competences and qualification routes in Europe*. Luxembourg: Office for Official Publications of the European Communities, 2009
- Coulet, J.C. (2011). La notion de compétence : un modèle pour décrire, évaluer et développer les compétences. *Le Travail Humain*, 74 (1), 1-30.
- Department of Education and Skills (2016) *Programme Recognition Framework: Guidance Counselling*. Dublin, Ireland: DES
- European Lifelong Guidance Policy Network (ELGPN) (2015) *Guidelines for Policies and Systems Development for Lifelong Guidance: a reference framework for the EU and for the Commission*. Jyvaskyla: ELGPN
- Galesic, M., & Bosnjak, M. (1999). Effects of questionnaire length on participation and indicators of response quality in a web survey. *Public Opinion Quarterly*, 73, 349–360.
- Herr, E.L., Cramer, S.H., and Niles, S.G. (2004) *Career guidance and counselling through the lifespan* (6th ed.) Boston: Pearson, Allyn and Bacon
- International Test Commission. (2006). International Guidelines on Computer-Based and Internet- Delivered Testing, *International Journal of Testing*, 6 (2), 143-171.
- McCarthy, J. (2002) *The training, skills and qualifications of career guidance workers* (OECD expert paper). Paris: OECD
- McMahon, M. (2019) Qualitative career assessment: a higher profile in the twenty-first century? In Athanasou, J. A., and Perera, H.N (eds.) *International Handbook of Career Guidance*. Switzerland: Springer Nature
- McMahon, M., Watson, M., and Lee, M.C.Y (2019) Qualitative career assessment: a review and reconsideration. *Journal of Vocational Behaviour* 110, p. 420-432
- McMahon, M., Patton, W., and Watson, M. (2003) Developing qualitative career development processes. *The Career Development Quarterly* 51, p.194-202

McMahon, M., Watson, M., and Patton, W. (2005) Qualitative career assessment: developing the My System of Career Influences Reflection Activity. *Journal of Career Assessment* 13 (4) p. 476-490

National Career Development Association (2015) *Code of Ethics*. Broken Arrow, OK: NCDA

Stoltz, K.B., and Barclay, S.R (Eds.) (2019) *A comprehensive guide to career assessments*. (7th ed.) Broken Arrow, OK: National Career Development Association

Niles, S. G., & Harris-Bowlsbey, J. (2021). *Career development interventions* (6th ed.)

Columbus, OH: Pearson Publishing.

Osborn, D.S. and Zunker, V.G. (2016) *Using assessment results for career development* (9th ed.) Boston: Cengage Learning

Osterlind, S.J. (2010) *Modern measurement: theory, principles, and applications of mental appraisal* (2nd ed.) Boston: Pearson, Allyn and Bacon

ANNEX ONE

EUROPEAN UNION AND NATIONS

TABLE OF FREE ONLINE TESTS PROVIDED ON
GOVERNMENT SUPPORTED WEBSITES

BASED ON FOLLOW-UP TO THE EUROPEAN COUNTRIES AND NATION
RESPONSES ON ICT IN THE CEDEFOP LIFELONG GUIDANCE SYSTEMS AND
PRACTICES INVENTORYⁱ 2019-20

COUNTRY	MINISTRY /AGENCY RESPONSIBLE	SUBJECT OF TEST OR QUESTIONNAIRE	FUNCTION OF TEST OR QUESTIONNAIRE	PSYCHOMETRIC /NON-PSYCHOMETRIC	ACCOMPANIED /SELF-SERVICE
AUSTRIA	AMS (Public Employment Service) www.berufsscompass.at	Occupational Interests (Holland's theory); Personality (Big Five theory), Personal strengths, Job expectations	-Self-appraisal -Career exploration	Psychometric Psychometric	Self-service
	Institute for Economic Research in Education	Interests	-Self-appraisal		

	https://www.bic.at/bic_interessenprofil.php		-Career exploration	Non- psychometric	
BELGIUM (Fl.)	Ministry of Education and Training, Flanders https://www.onderwijskiezer.be/	Interests, Study attitude, Motivation, and method,	-Self-appraisal -Career exploration	Non- psychometric	Accompanied offline
BULGARIA	Ministry of Education and Science https://orientiranemon.bg/	Occupational interests, abilities, preferences, values	-Self-appraisal -Career exploration	Non- psychometric	Self-service
CROATIA	Croatia Science Foundation – Ministry of Education https://www.putkarijere.hr/upitnici/upitnici	Interests (Holland’s theory; Personal Globe) Values, life roles, and career maturity (Super), Career adaptability, and vocational identity (Savickas),	-Self-appraisal -Career exploration -Decision making problem identification	Psychometric	Self-service

		Decision-making (Gati), Occupational and learning pathway awareness			
CYPRUS	No free publicly funded online tests				
DEN-MARK	Ministry of Education www.ug.dk	Strengths, interests	-Self-appraisal	Non-psychometric	Accompanied online
ESTONIA	Ministry of Labour-Estonian Unemployment Insurance Fund: partner company https://www.tripod.ee/?keel=inglise	Interests (Holland's theory), Personality (Big Five), Mental Ability	-Self-appraisal -Career exploration	Psychometric	Accompanied offline
FINLAND	Ministry of Labour www.asiointi.mol.fi/	Interests (Holland's theory)	-Self-appraisal -Career exploration and learning	Non-psychometric	Accompanied offline Smart phone version
	Universities of Applied Science and Vocational Colleges www.crear.fi	Study, well-being, resilience, employment and career planning skills	Self-appraisal	Non-psychometric	Accompanied offline if requested by test-taker
FRANCE	ONISEP (Ministry of Education) www.onisep.fr	-Interests	-Career exploration	Non-psychometric	Accompanied online

	Ministry of Education https://www.parcoursup.fr/index.php?desc=services_numeriques	Higher education interests and choice	Career exploration and learning	Non-psychometric	Accompanied offline
	Ministry of Labour: Pole Emploi (Public Employment Service) www.emploi-store.fr	-Interests (Holland's theory) -Values, interests, and motivation for work	-Self-appraisal -Career exploration	Psychometric	Accompanied offline Smart phone version
	Ministry of Labour https://app.diagoriente.beta.gouv.fr/	-Transversal competence experiences (Bandura's Social Learning; Coulet on defining individual and collective competences) (based on multi-disciplinary expertise, EU cooperation, and national field testing), Interests, Possible pathways, job search	-Self-appraisal -Transversal competence identification -Career exploration and job tasting	Psychometric Non-psychometric	Accompanied offline
GERMANY	Federal Employment Agency (Federal	Skills, Social skills, Interests,	-Self-appraisal	Psychometric	Self-service with offline

	Ministry of Employment)	Occupational preferences (Holland theory, extended to 8 categories)	-Career exploration		accompaniment
	https://www.arbeitsagentur.de/bildung/welche-ausbildung-welches-studium-passt				
GREECE	Ministry of Education, Research, and Religious Affairs: National Organisation for the Certification of Qualifications and Vocational Guidance www.eoppep.gr	Interests (Holland's theory), Work values, Decision-making ability	-Self-appraisal -career exploration	Psychometric	Self-service with offline accompaniment
HUNGARY	Ministry of National Economy https://palyaorientacio.munka.hu/	Skills, interests	-Self-appraisal	Non-psychometric	Self-service
IRELAND	Department of Education and Skills: Quality and Qualifications Ireland www.qualifax.ie	Interests (quasi-Holland)	-Self-appraisal -careers exploration	Psychometric	Self-service
ITALY	Province of Varese	Competences, values,	-Self-appraisal	Non-psychometric	Self-service

	www.mioriento.it	planful, personal qualities			
LUXEM- BOURG	Ministry of Education, Children, and Youth: Maison de l'orientation www.anelo.lu	Interests, skills, personal qualities	-Self- appraisal -careers exploration	Psychometri c Non- psychometri c	Self- service Offline support offered
MALTA	No free online publicly funded tests				
NETHER -LANDS	Ministry of Education, Culture, and Science www.kiesmbo.nl www.studiekeuze1 23.nl <a href="https://www.lerene
nwerken.nl/aan-de-
slag-met-je-
loopbaan">https://www.lerene nwerken.nl/aan-de- slag-met-je- loopbaan	-Interests -career planning -personality (Holland's theory)	-Self- appraisal -careers exploration	Non- psychometri c	Self- service
NORWAY	Ministry of Education- Skills Norway <a href="https://utdanning.n
o/karriereplanlegge
ren/">https://utdanning.n o/karriereplanlegge ren/	-Career planning skills	-Self- appraisal	Non- psychometri c	Self- service with online and offline support offered
PORTU- GAL	Institute for Employment and Vocational Training (public	-personality, -interests, -values,	-Self- appraisal -careers exploration	Non- psychometri c	Self- service

	employment service) www.iefponline.iefp.pt	-abilities, -skills			
	Ministry of Education	-Personal qualities, -Knowledge of education and training pathways, occupations, and working conditions	-Self-appraisal -careers exploration	Non-psychometric	Self-service
ROMANIA	No free publicly funded online tests				
SLOVENIA	Employment Service of Slovenia https://esvetovanje.ess.gov.si/	Personality, (Myers Briggs approach) Interests (Holland's theory), values, motives, work style, competency, decision-making capacity, personality (4 tests)	-Self-appraisal -careers exploration	Non-psychometric Psychometric	Self-service - Accompanied during test taking

SPAIN	Ministry of Education and Vocational Training https://www.todofp.es/orientacion-profesional.html	Learning and work skills, interests	-Self-appraisal -careers exploration	Non-psychometric	Self-service
SWEDEN	Swedish Public Employment Service https://arbetsformedlingen.se/for-arbetssokande/yrken-och-framtid/intresseguide/	Interests (Holland's theory)	-Self-appraisal -careers exploration	Non-psychometric	Self-service
UK NATION ENGLAND	Department of Education and Skills National Careers Service https://beta.nationalcareers.service.gov.uk/	Interests Motivations Preferences Skills	-Self-appraisal -careers exploration	Psychometric	Self-service
UK NATION SCOTLAND	Department of Education Skills Development Scotland https://www.myworldofwork.co.uk/	-Personality (Holland's theory) -Strengths (Myers Brigg theory) -Skills	-Self-appraisal -careers exploration	Non-psychometric Psychometric	-Self-service -Results can be shared with practitioner during career counselling

¹ <https://www.cedefop.europa.eu/en/publications-and-resources/country-reports/inventory-lifelong-guidance-systems-and-practices>

Twenty-one Member States of the European Union, one EEA country (Norway), and two UK nations provided information for the CEDEFOP inventory. The six EU countries not covered in the inventory are Croatia, Czech Republic, Latvia, Lithuania, Poland, and Slovakia. Croatia provided additional information specifically for this SA-EU research task.

ANNEX TWO

METHODOLOGICAL NOTE ON REVIEW OF EUROPEAN UNION WEBSITE TESTS

The Terms of Reference (TOR) for this desk research required the EU desktop report to address the following:

- a. Determine EU member states with existing guidelines and standards for non-psychometric assessments
- b. Investigate the legislative framework for managing psychometric and non-psychometric assessments in each member state
- c. Study best practice in developing, applying, and managing non-psychometric career assessment guidelines in five EU countries.

The criteria for selecting EU countries were as follows:

- i) A mix of countries that have been implementing online psychometric and non-psychometric assessment for 1 to 5 years, and 5 years or more
- ii) Countries that have a statutory professional body separate from the government entity/agency that implements online assessments
- iii) Countries that have assessed or can report on the efficacy and uptake of the online assessments
- iv) Countries that represent both the former Eastern Bloc and Western Europe.
- v) Countries that DHET has an interest in: Hungary, Luxembourg, and Ireland.

With reference to a) above, the international literature review (Section 2), undertaken as a first step, showed that formal guidelines and standards for non-psychometric tests do not exist anywhere in the world. Thus, it would be impossible to provide answers to a) and likewise to c). With reference to b), it would also be impossible to find answers concerning the legislative environment for non-psychometric tests. Answers concerning the legislative environment for psychometric tests are tempered by the findings of international literature review (Section 1) which clearly show that standards for the development of psychometric tests are guidelines for best practice. They are not legally enforceable nor is their implementation monitored and controlled by the professional associations. Tests (and test standards) are also viewed differently in terms of whether the test results have high stakes or low stakes consequences. The AERA, APA, and NCME *Standards* (2014) are the gold standard guidelines for psychometric test development, administration and score interpretation worldwide, and are internationally recognized, including within the European Union.

The literature review also highlighted separate and related issues for non-psychometric test development and usage: online testing, ethics, and the training of career practitioners, thus providing a broader context for viewing the research task than foreseen in the TOR and for addressing the basic research questions.

As regards the choice of five EU countries, bearing in mind the points just raised, it became clear from initial inquiries that while there was an overview of ICT for career learning in 19 of the 27 EU member states and in one EEA state (CEDEFOP, 2020), there was no EU overview of the use of free online publicly funded career assessment tests. Thus, the use of the TOR criteria, i) to v) above, for choosing five EU countries, could result in having a very unrepresentative sample that would shed little light on answers to the basic research questions. For all of the reasons given above, it was deemed that the methodology of developing five country case studies would not produce meaningful answers to the research questions.

For these reasons, a broader picture of practice across many European countries was sought. The CEDEFOP inventory provided a starting point. From this data, online tests in 21 countries and nations were researched, identified and reviewed and contact was made with national correspondents of CEDEFOP's CareersNet and other national experts in seeking more information about the tests. Further information was sought from Croatia. Three countries have no free publicly funded online tests. This different methodological approach (to that of the TOR) provides a wider panorama of examples of free online publicly funded career assessment tests in the EU (see table in Annex 2) that DHET can draw upon for inspirations for future online non-psychometric test development and provision in South Africa, as well as names of contact persons in different countries for organizational and professional exchange (Annex 4). The panorama of practice in 22 countries is more representative of the international and EU context for free publicly funded online career assessments than any five EU country case study sample could give.

ANNEX THREE

LIST OF PERSONS COMMUNICATED WITH FROM WHOM ASSISTANCE WAS RECEIVED FOR THIS REPORT

COUNTRY/ INTER- NATIONAL ORGAN- ISATION	NAME	AFFILIATION AND ROLE	ROLE IN THIS RESEARCH TASK	CONTACT DETAILS
AUSTRIA	Wolfgang Bliem	Institute for Economic Research in Education (IBW) Vienna	Key national informant	bliem@ibw.at
	Judith Csarmann	Labour Market Research and Careers Information Department, Federal Office, Employment Service Austria, Vienna	Key test informant	judith.csarmann@ams.at
	Martina Ronnenberg	(do.)	Key test informant	martina.ronnenberg@ams.at
BELGIUM (Fl.)	Claire Kagan	CEDEFOP CareersNet representative	General informant	
BULGARIA	Lachezar Afrikanov	CEDEFOP CareersNet representative	General informant	
CROATIA	Irena Bacelic	Head of Labour Market Section, Ministry of Labour, Zagreb	Key national informant	irena.bacelic@mrm.s.hr
DENMARK	Steffen Jensen	Retired Head of Unit, Ministry of Education and Children, Copenhagen	Key national informant	sj@steffenjensen.dk
ESTONIA	Kristina Orion	Deputy Head, Department of Skills and Career Services, Estonian	Key national informant	Kristina.Orion@tootukassa.ee

	Margit Rammo	Unemployment Insurance Fund, Tallinn Estonian Agency for Erasmus+ and for the Solidarity Corps, Education And Youth Board, Tallinn	Key national informant	Margit.Rammo@harno.ee
FINLAND	Dr Raimo Vuorinen	Lecturer, Institute of Educational Research, Jyvaskyla	Key national informant	Raimo.vuorinen@jyu.fi
FRANCE	Graziana Boscato	National Euroguidance Coordinator, Rectorat, Academie de Strasbourg	Key national informant	graziana.boscato@ac-strasbourg.fr
	Pascal Chaumette	Director D6 – Diagoriente, Paris	Key test informant	pascal.chaumette@id6tm.org
	Fabien Beltrame	Test developer. Inspector General, Direction Générale de Pôle Emploi / DIGAI, Paris	Key test informant	fabien.beltrame@pole-emploi.fr
GERMANY	Sascha Zirra	Test developer, psychometrist. Head of Unit, Media for Occupational Orientation and Occupational Information, Federal Employment Service, Nurnberg	Key test informant	Sascha.Zirra@arbeitsagentur.de
	Kathrin Schoenle	Expert Product Owner (do.)	Key test informant	Kathrin.Schoenleben@arbeitsagentur.de

GREECE	Fotini Vlachaki	Lifelong Guidance Expert - Researcher Ministry of Labour and Social Affairs, Athens	Key national informant	fvlachaki@yeka.gr
HUNGARY	Dr Tibor Bors Borbely Pezce	Ministry of Economy	Key national informant	borsborbely@yahoo.com
IRELAND	Jennifer McKenzie	Director, National Centre for Guidance in Education.	General informant	Jennifer.mckenzie@ncge.ie
	John Carton	Test developer, website owner: The Careers Portal	Key test informant	jcarton@careersportal.ie
ITALY	Giulio Iannis	CEDEFOP CareersNet representative	General informant	
LUXEMBOURG	Jean Jacques Ruppert	Head of Unit, AVOPP, Luxembourg	Key national informant	jeanjacquesruppert@me.com
	Jos Noesen	Retired Ministry of Education, Luxembourg	General informant	jos.noesen@education.lu
NETHERLANDS	Karien Coppens	Senior Researcher, Centre for Vocational Education and Training, Randstad	Key national informant	Karien.Coppens@ecbo.nl
NORWAY	Ingjerd Espolin Gaarder	Department Director, Skills Norway, Oslo	Key national informant	ingjerd.gaarder@kompetansenorge.no
PORTUGAL	Helia Moura	Senior Advisor, Directorate General for Education, Ministry of Education, Lisbon	Key national informant	helia.moura@dge.mec.pt
SLOVENIA	Miha Lovsin	Head of Department, Office for the Development and Quality of Education, Ministry of Education and Sport, Ljubljana.	Key national informant	Miha.Lovsin@gov.si
	Stasa Bucar Markic	Senior Field Consultant, Service	Key test informant	Stasa.BucarMarkic@ess.gov.si

		Development for Lifelong Guidance, Public Employment Services, Ljubljana.		
SPAIN	Luis Carro	CEDEFOP CareersNet representative	General informant	
SWEDEN	Nina Ahlroos	Senior Advisor, Swedish Council for Higher Education	Key national informant	nina.ahlroos@uhr.se
UK-ENGLAND	Siobhan Neary	Associate Professor and Head of the International Centre for Guidance Studies (iCeGS) Institute of Education University of Derby.	Key national informant	s.neary@derby.ac.uk
	Ray Plummer	National Careers Service Team	Key test informant	Ray.PLUMMER@education.gov.uk
		Education and Skills Funding Agency		
UK-SCOTLAND	Derek Hawthorne	Digital Development Manager, Skills Development Scotland.	Key test informant	Derek.hawthorne@sds.co.uk
	Sandra Cheyne	National CIAG Policy & Professional Practice Lead, Skills Development Scotland	General informant	Sandra.cheyne@sds.co.uk
USA	Prof. Spencer Niles	Theorist, researcher, test developer, author William and Mary University, VA	Key international informant	sgniles@wm.edu
	Dr Kevin Stoltz	Researcher, author, National Career Development Association (NCDA), University of North Alabama.	Key international informant	kstoltz1@una.edu
	Pam Frugoli	O*NET and Competency Model Team Lead,	Key test informant	Frugoli.pam@dol.gov

		Employment and Training Administration, U.S. Department of Labor		
AUSTRALIA	Prof. Mary McMahon	Theorist, researcher, test developer, author Honorary Associate Professor School of Education The University of Queensland Brisbane	Key test informant	marylmcMahon@uq.edu.au
	Prof. James Athanassou	Researcher, test developer, psychometrist, author. Associate Professor [Honorary] Faculty of Health Sciences, Discipline of Rehabilitation Counselling The University of Sydney	Key test informant	james.athanassou@sydney.edu.au
CEDEFOP	Cynthia Harrison Villalba	Researcher, project manager, CEDEFOP, Thessaloniki, Greece	General informant	Cynthia.HARRISON-VILLALBA@cedefop.europa.eu
MALTA	Dorianne Gravina	Education Officer (Career Guidance) National School Support Services(NSSS) Ministry of Education and Employment	Key national informant	dorianne.gravina@educ.gov.mt