

# Testing Competences Worldwide in Large Numbers

## Complying with ISO/IEC 17024:2012 Standard for Competence Testing

Marianne Hubregtse<sup>1(✉)</sup>, Stefanie Moerbeek<sup>1</sup>, Bernard P. Veldkamp<sup>2</sup>, and Theo Eggen<sup>3</sup>

<sup>1</sup> EXIN Holding B.V., Utrecht, The Netherlands  
{marianne.hubregtse, stefanie.moerbeek}@exin.com

<sup>2</sup> University of Twente, Enschede, The Netherlands  
b.p.veldkamp@utwente.nl

<sup>3</sup> Cito B.V., Arnhem, The Netherlands  
theo.eggen@cito.nl

**Abstract.** The ISO/IEC 17024:2012 describes best-practices regarding competence assessment. The standard emphasizes validity and sound processes to create high-quality competence assessments. The question is how to comply with the ISO/IEC standard when creating large-scale, worldwide assessments. In addition, the competence framework describes competences that require years of experience as part of the competence. We determine to what extent of mastery candidates need to master the competence to start working. We assess this by testing the requisite knowledge with a multiple-choice exam and the required minimum level of mastery of the competence with a Practical Assignment. This assignment is assessed by trainers, which creates the need for supervision and accreditation of the trainers. This paper shows an example of a certification scheme to explain how we comply with the ISO/IEC standard. The creation of the certification scheme and the accreditation of the trainers are described. The compliance with the ISO/IEC standard is explained.

**Keywords:** Competence assessment · ISO/IEC 17024:2012 · Assessment quality · Large-scale assessment · Worldwide assessment · Validity · e-Competence framework

## 1 Introduction

### 1.1 Testing Competences

Competences are a combination of knowledge, skills and attitudes [1, 2]. Testing *knowledge* can be done fairly easily, by using multiple-choice exams. Testing *competences* is more complex.

Proving that a candidate masters a competence is difficult. Some competence frameworks describe a competence by defining the knowledge, skills and attitude of a professional working in that area for a longer period of time. The e-CF framework for ICT competences [3] that we work with is one of those frameworks. Even though the competences are described for professionals with at least 5 years of experience, a professional

starting out in a job is not completely void of any competence; rather they master part of the competence as described in the framework. Employers still want to know in which competences a professional has started working towards complete mastery. Therefore, exams can be designed in a way that allow candidates to show their partial mastery.

Sometimes, this means that the exams need to give the candidates the opportunity to show that they have the requisite background knowledge. In that case, we can show the partial mastery with a multiple-choice exam. In other cases, a professional needs to show that they master basic tasks, have the requisite background knowledge and have an adequate attitude, to work towards full mastery of the competence. Consequently, some certification schemes will incorporate only multiple-choice exams, while others will have a mix of both multiple-choice exams and practical assignments.

For example, when creating an exam for a starting Scrum Master, it must be determined to what extent a described competence must be shown by a starter. Suppose that a relevant competence states that the candidate: “Takes proactive action and develops organizational processes to address the development needs of individuals, teams and the entire workforce.” (Manage, D.9. Personnel Development, level 4) [3]. It might be relevant for employers to attract a starting Scrum Master that has shown that she knows how to take action to help individual team members to develop their competences. The proactivity and the team needs could only be relevant for professionals with a few years of experience.

When testing partial mastery of competences, the test goal and exam specifications must specify exactly which competences are tested to exactly what extent. In addition, it must be explained why the experts judge the partial mastery to be enough for starting professionals. In some cases, it might be that showing the requisite knowledge is enough to start working. We take a pragmatists view to competence testing in this case.

For instance, suppose that we want to create an exam that shows that a Software Developer can start developing in PHP. It is judged enough to test whether the candidate has the necessary knowledge. Suppose that a relevant competence is: “Acts under guidance to develop, test and document applications.” (Build, B.1. Application Development, level 1) [3]. The same competence asks for the following knowledge components: “The candidate is familiar with

- the appropriate software programs/modules;
- hardware components, tools and hardware architectures;
- [etc....]” (Build, B.1. Application Development, Dimension 4, 1<sup>st</sup> 2 knowledge examples) [3]

In this case it is reasonable to test whether the candidate knows enough PHP to start coding under guidance of a more experienced coder.

## 1.2 Worldwide Examination

When assessing competences all over the world, a few challenges are introduced.

Firstly, the validity of the certification scheme must be proven in an international context. We do this by collaborating with international subject matter experts and training providers, but it stays important to test localizations for validity within that local context.

Secondly, when competences are tested, training and practical assignments are always part of the certification scheme. However, trainers and supervisors are not EXIN employees, so that we can maintain impartiality. As a certification body, we are responsible for assessing and scoring. By allowing the trainers and supervisors to act as assessors, we outsource part of our work. The challenge lies in ensuring that competent and honest trainers and supervisors do the assessments.

### 1.3 Large-Scale Assessments

We sell around 150,000 exams per year. Most of these exams are exams that only test the requisite knowledge of candidates to start their professional careers. These are multiple-choice exams. About 64 % of these exams are taken online using Online Proctoring. The candidate is required to log onto a secure web-environment, show proof of identification and to allow that for each exam sound and video are recorded. All videos are looked at fully (on high-speed) by an employee to signal any indication of fraud.

The other 36 % are paper-based exams. These exams have bubble-sheets that allow for automated scoring of the exams. The forms are read in by scanner and email, or by mobile application through a photograph. Only supervisors are allowed to send in the exam forms.

The scoring of practical assignments cannot be automated, because the criteria always need to be interpreted by an expert. However, we are in the process of allowing trainers or assessors to directly input their scoring of the individual candidates into our database. As of the writing of this paper, this is not fully done yet. Currently, supervisors and trainers score the candidates on paper, using the provided observation criteria, and only report the result (passed/failed) to our database. Changing this will involve training the supervisors and trainers to use the system correctly.

## 2 The ISO/IEC 17024:2012 Standard

The ISO/IEC 17024:2012 standard [2] describes best practices regarding assessing competences. As an exam institute, it is important for us to be ISO/IEC 17024:2012 certified. Not only does certification give us more credibility as an exam institute issuing certificates, we are also genuinely concerned with quality and think that the ISO/IEC standard reflects best practices.

In this paper we will define *certification scheme* as it is described in the ISO/IEC standard (Article 3.2) [2]: the “competence and other requirements related to specific occupational or skilled categories of persons”. In a certification scheme, all requirements, that a candidate needs to fulfill before obtaining a certificate, are described.

For the purposes of clarity, *exam* is here defined as a multiple-choice test of the knowledge of candidates on a certain topic. *Assessment*, is defined as any test that allows a candidate to demonstrate their extent of mastery of a competence.

The ISO/IEC norms do not always specify exactly which processes you need to follow. While getting certified, additional questions were asked about the processes of creating the certification scheme, maintaining impartiality whilst working with a select few experts and ensuring high validity exams.

We will shortly discuss the ISO/IEC 17024:2012 papers that elicited questions, to show what auditors asked after most (in our case). A short definition of validity, that we can agree with, is given as well, since the ISO/IEC standard does not define the concept of validity.

The NEN-EN-ISO/IEC 17024:2012 standard [2] is a best-practices guideline when creating the certification scheme. However, although it gives recommendations on what elements regarding the certification scheme must be described (ISO 17024:2012, Paper 8) [2], it does not give practical guidelines or examples.

Specifically, there is no information in the standard on creating large-scale and worldwide competence assessments, nor on how to deal with testing partial mastery of competences. For us, it was difficult to justify exactly how we complied with the ISO/IEC standard, even though we felt we were on the right track. We show how one might argue this compliance. We show a real example of a certification scheme used for worldwide competence assessment. We also describe how we have solved the issue of taking full responsibility for the quality of the assessment, while outsourcing the assessment.

## 2.1 Compliance with ISO/IEC 17024:2012

The ISO/IEC 17024:2012 standard [2] is quite extensive and, therefore, this paper does not cover all parts of it. However, three of the papers are relevant here: Paper 8 regarding the validity of the certification scheme, Paper 5 regarding the impartiality of the certification body and Paper 6 regarding outsourcing work.

**Validity.** ISO/IEC 17024:2012 states that:

“A certification scheme shall contain the following elements:

- (a) scope of certification;
- (b) job and task description;
- (c) required competence;
- (d) abilities (when applicable);
- (e) prerequisites (when applicable);
- (f) code of conduct (when applicable).” (Paper 8.2) [2]

These elements in the certification scheme help build a validity argument (see also Sect. 2.1). These instructions also underline the importance of specifying the partial competences tested, complete with a specification of the knowledge and skill elements that need to be mastered (and thus tested). The example will show one way of building these elements.

We define a valid exam as an exam where the score yields information that you can use to make decisions about a candidate [4, 5]. Validity can be made plausible by showing the link between all tasks the candidate may encounter as a professional, the competences necessary for performing those tasks and the chosen questions and assignments for the certification scheme [5, 6] as described in the test goal and the exam specifications.

We ensure validity of the certification schemes by allowing professionals (often subject matter experts) and trainers to be involved in the creation of the scope, the job and task description, the required competences and abilities. In addition, we discuss the prerequisites and codes of conducts for certification with subject matter experts.

**Impartiality.** In order for candidates to be assessed fairly, it is often recommended to separate the training provider and the certification body [7–17]. EXIN is an independent certification body and does not provide training to candidates, which makes it easy to comply with the ISO/IEC 17024:2012 (Paper 5.2), which states that “Offering training and certification for persons within the same legal entity constitutes a threat to impartiality [2].”

Since we do not train candidates, there is absolute impartiality; we do not benefit from candidates passing or failing assessments. However, not training candidates also means that there are no opportunities for our employees to come directly into contact with candidates during their performances for their assessments. This means that we need professionals and trainers to help us ensure validity, and thus we need to outsource part of the work on a certification scheme.

**Outsourcing Work.** According to the ISO/IEC 17024:2012 standard, the certification body remains responsible for the outsourced work and thus must “ensure that the body conducting outsourced work is competent and complies with the applicable provisions of this International Standard;” (Paper 6.3.2.b) [2]. A challenge lies in ensuring that assessors that you rely upon to assess a candidate’s performance are competent and comply with the rules you have set for the assessment.

We ensure that assessors comply with our rules by accrediting our training providers, training the trainers and auditing them on a regular basis. In order to ensure competence of the assessors, we ask for work experience in a relevant area, references that confirm the work-experience and the successful completion of the exam that they will assess. Trainers with ample work experience in a relevant area are exempt from the mandatory training and practical assessment. Other candidates never are.

### 3 Example: EXIN Agile Scrum Master

This part of the paper will give a description of the processes used to create the certification scheme for Agile Scrum Master. Please note that the certification scheme was still in development whilst writing this paper. All examples given are subject to change during the development, but they do reflect an example of what the final product could look like. The processes followed to create this certification scheme give insight in how we deal with maintaining impartiality, ensuring validity and outsourcing work regarding the creation of the certification scheme and the assessment of the candidates.

#### 3.1 Certification Scheme

The Preparation Guide is our central documentation of the certification scheme. This guide is freely available to anyone on our website. It contains all elements for a

certification scheme as listed by the ISO/IEC 17024:2012 standard. Candidates and trainers can refer to this document to prepare for the assessments.

**Scope of Certification.** Agile Scrum is a project management method for software development. A small team (3–9 people) works in short iterations of time, to deliver new functionality. Every new iteration, the list of requirements for the software is updated and prioritized. This creates great flexibility for the customer.

Scrum knows three major roles: Product Owner, Scrum Master and Development Team member. The team is self-managing. Therefore, there is no need for a traditional project manager. Instead, the Product Owner is the voice of the customer and helps prioritize features for the next iteration. The Scrum Master coaches the team to be self-managing through servant leadership and training. A Scrum Master also keeps track of the progress of the project.

**Process.** The scope of the certification is determined based on market research. A survey was sent out under 54 partners (mostly training companies) and candidates, to generate the scope of the exam. Allowing our partners and candidates to give input on new certification schemes shows the market value and adds to the validity.

**Test Goal.** The goal of the Agile Scrum Master certification scheme is to gather enough information on the competences of the candidate to determine whether a candidate is ready to perform the desired tasks to the desired level, and thus deserves a certificate. This means that a candidate with an Agile Scrum Master certificate must be able to function in the role of Scrum Master. The candidate is not expected to master the competences as a professional with a few years of experience would. Rather, the candidate must show that she has just enough competence to start working as a Scrum Master for the first time. In addition, the candidate must show that she has the requisite knowledge to perform the function.

**Job and Task Description.** In this case, there is already a solid framework that describes full competences: European e-Competence Framework (e-CF) [3]. Instead of defining our own competences for every single certification scheme, all EXIN certification schemes use the e-CF as a common framework. The complete e-CF represents the practice domain of the Agile Scrum Master. The selection of the competences represents the competence domain.

**European e-Competence Framework.** The practice domain is described in the e-CF [3]. The e-CF describes ICT competences in the five main areas Plan, Build, Run, Enable and Manage. The levels within each competence give an indication of the level of responsibility that is required: a higher level indicates more responsibility. In essence, the e-CF is a job and task analysis for the five main areas; professionals in the ICT work field collaborated to create the e-CF.

**Process.** EXIN employs exam experts, that are trained on best practices in assessment and exam creation. Since the certification portfolio covers a broad part of the ICT work field, we rely on subject matter experts for the content of the questions and practical

assignments. Since we certify candidates worldwide, we work together with subject matter experts from all over the world. We use online authoring methods to work together on content and questions. The content of the job and task description is supplied by subject matter experts. We select the subject matter experts on the basis of their demonstrated or verified experience and earned certificates.

For Agile Scrum Master, two exam experts guided two international subject matter experts in building the job task analysis from the e-CF. Firstly, both subject matter experts individually selected the relevant competences from the framework. Then an online video conference was held under supervision of the exam experts, where the subject matter experts agreed on the relevant competences and level. The result is shown in Table 1.

**Table 1.** Example e-CF mapping for agile scrum master

Area	Competence name	e-Level	Extent
Plan	A.2. Service Level Management	4	Superficial
Plan	A.5. Architecture Design	3	Superficial
Build	B.2. Component Integration	3	Superficial
Build	B.3. Testing	3	Superficial
Enable	D.3. Education and Training Provision	3	Partial
Enable	D.9. Personnel Development	2	Partial
Manage	E.3. Risk Management	3	Partial
Manage	E.5. Process Improvement	3	Partial

**Mastery of Competences.** The e-CF mapping alone is not enough to start developing the assessment. In addition to the mapping, it must be decided to what extent the competence level should be represented in the certification scheme. Furthermore, it should be decided which knowledge components a candidate should be able to show.

**Process.** The extent to which a candidate must show mastery of a competence could fall into one of the following categories: general, partial or superficial. For each competence in the e-CF, we have developed a set of observation criteria. The full set of criteria for each competence is extensive, but when a candidate has shown that they can perform all tasks listed in the criteria (as a professional with experience often can), they are awarded credit for the full competence.

The extent of the mastery that is tested within a certain certification scheme is decided by the number of observation criteria that are assessed through the practical assignments. When 1 % to 29 % of the total number of criteria are assessed, the competence is regarded as covered superficially. Between 30 % and 69 % coverage of the criteria is regarded as partial. When 70 % or more of the criteria of a competence level are covered by the certification scheme, we regard the competence as generally covered. The subject matter

experts, under guidance of the exam experts, decided which observation criteria are relevant for the scope of the exam and the test goal.

After this was decided, the subject matter experts and the exam experts agreed on the relevant knowledge components and translated these to the exam requirements, which form the basis for developing questions for the multiple choice exam.

**Assessment Process.** The Agile Scrum certification scheme consists of a multiple-choice exam, a mandatory training and successful completion of the Practical Assignments. The trainers are responsible for assessing whether the candidate has shown adequate competence, in the Practical Assignments. What is ‘adequate’ is determined by the chosen observation criteria. The trainers are provided with material and observation criteria that show under which conditions a candidate is eligible for successful completion of the Practical Assignments. The trainers must use the Practical Assignments issued, but they may adapt to their context, in order to allow candidates to show their mastery of the competences. Where possible, the assignments have a clear rating scale, to help the trainer assess.

As mentioned earlier, we keep control over the assessment by accrediting the trainers. In addition, the required multiple-choice exam ensures that we directly control at least half of the scoring of the assessment.

**Skills and Attitude Assessment.** As can be seen from Table 1, Agile Scrum Master does not cover any of the relevant competences generally. It is important to specify which observation criteria are seen as relevant for Agile Scrum Master, so that all trainers may assess candidates as uniformly as possible.

The observation criteria for the competence D.3. Education and Training Provision (level 3) are: The candidate can...

- address organizational skills needs and gaps
- *adapt training plans to address changing demands*
- promote and market training provision
- design curricula and training programs
- *establish a feedback mechanism*
- *implement continuous improvement of education and training delivery*
- *assess added value of training provision*

Of these criteria, the italicized criteria were chosen by the subject matter experts as relevant for a starting Agile Scrum Master. These are 4 out of 7 criteria, or 57 %, so we call the competence covered partially. The same process was repeated for all other competences in this certification scheme.

**Knowledge Assessment.** After determining the competences relevant for the certification scheme, we asked the subject matter expert to identify all the requisite knowledge for a starting Scrum Master. This list of requisite knowledge is captured in the exam blueprint, which is made available for trainers and candidates in the Preparation Guide. The resulting exam blueprint is shown in Table 2. The Agile Scrum Master exam consists of 40 multiple-choice questions, divided over the exam requirements. The questions



allow the candidate to show that she possesses the requisite knowledge to start as a first-time Scrum Master.

**Table 2.** Example exam blueprint for agile scrum master

Exam requirements			# Questions
<b>1. Agile Way of Thinking</b>			
	1.1	Agile concepts	2
	1.2	Continuously improving the process	1
	1.3	Other Frameworks and other Agile frameworks	2
	1.4	Applying Agile principles to IT Service Managements	1
<b>2. Scrum Master Role</b>			
	2.1	Responsibilities and Commitment	3
	2.2	Coaching the Team and Mediating	3
	2.3	Other roles (Product Owner, Development Team)	3
<b>3. Agile Estimating, Planning, Monitoring and Control</b>			
	3.1	Writing and maintaining the Product and Sprint Backlog	3
	3.2	Agile Planning	2
	3.3	Agile Estimation	4
	3.4	Tracking and communicating progress	3
	3.5	Staying in control	1
<b>4. Complex Projects</b>			
	4.1	Scaling Agile Projects	2
	4.2	Suitability of Agile for different types of projects	2
	4.3	Agile administration in tooling and tool integration	1
<b>5. Adopting Agile</b>			
	5.1	Introducing Agile	3
	5.2	Self-organization	2
	5.3	Agile requirements and proper environment	2
<b>Total</b>			<b>40</b>

The subject matter experts individually brainstormed about the requisite knowledge and then agreed in a video-conference on the final blueprint, under guidance of the exam experts. Subsequently, the exam requirements were worked out into exam specifications.

For example, the exam specifications for the exam requirement 2.3 Scrum Master are: The candidate can...

- identify which tasks are related to the role of Scrum Master.
- explain the competences required for performing the role of the Scrum Master.
- explain the tasks, responsibilities and authorities of the Scrum Master.

As can be seen from Table 2, the exam includes 3 questions for this exam requirement, so that each of the exam specifications can be represented in the exam. We try to assure that there is an equal number of questions and exam specifications, to ensure consistent exams.

When there are more exam specifications than questions in the exam, the subject matter experts are asked to agree beforehand on the exam specifications that are interchangeable in the exam. When there are fewer exam specifications than questions in the exam, the subject matter experts must agree on which specifications are represented by more than one question.

As soon as the exam specifications and requirements are accepted by both the subject matter experts and the exam experts, other international subject matter experts are asked to create the content for the actual multiple choice questions, under the guidance of the exam experts. The question creation process includes a review by both a subject matter expert and an exam expert, to ensure validity and quality. By asking different subject matter experts to determine and create the content of the assessment, we ensure international relevance and validity.

**Prerequisites and Code of Conduct.** The Preparation Guide includes all the prerequisites for the exam. In this case, all exams that show that the candidate understands the Agile Scrum framework were accepted as prerequisites. The code of conduct is not applicable for this certification scheme. (It would be applicable for a certification scheme Ethical Hacking, for instance.)

## 4 Discussion

We use the European e-Competence Framework (e-CF) [3] to create certification schemes. Since this framework describes competences that require work-experience, we test whether candidates have adequate mastery of the competence to start working with a Practical Assignment. Additionally, we test whether they possess the requisite knowledge to start working with a multiple-choice exam.

The example certification scheme for Agile Scrum Master complies with the ISO/IEC 17024:2012 standard for certification of competences. We have shown the processes we use to create all elements that need to be present in the certification scheme according to Article 8.2 [2]. We use subject matter experts to create valid, internationally relevant exams. To ensure the competence of the subject matter experts, we ask them to prove their experience and expertise to us, by means of work history and earned

certificates. In addition, we train the subject matter experts ourselves in item development best practices, using online training and self-study. In order to ensure impartiality, we always work with at least two experts, preferably from different countries, guided by at least one of our own exam experts.

We have shown how we determine the scope of the certification. By using both the e-CF and market research, we add validity and relevance to our exams. If we combine the scope of the exam with the input of at least two subject matter experts, we can create a very relevant, and thus valid, certification scheme. This combination is a form of the job and task description, as mentioned in the ISO/IEC standard.

We realize that there are other ways of conducting a job and task analysis. However, the main steps that need to be taken, are already undertaken in the creation of the e-CF: all competences have been described, complete with knowledge and skills examples. It seems efficient to use this information. We allow the subject matter experts to decide which competences of the e-CF enable the candidate to fulfill the scope of the certification. Nevertheless, we could have chosen to re-do that work, or to do the job and task analysis for all parts of the world separately.

We are also aware that adding more subject matter experts will change the outcome of discussions. In principle, the determination of the scope, as with any further work on the certification scheme, is not limited to two subject matter experts. In many cases, we work with more than two subject matter experts. However, we are bound by constraints of time and budget. We try to balance the implications of adding another expert to the development team with our constraints.

We have described a process of determining the extent to which we measure each competence. By using a fixed set of observation criteria for all competences to determine the extent, this is done in a repeatable way, making it less subjective and more comparable between certification schemes.

We realize that the fixed set of criteria does impose a limit. We could miss important criteria, by not allowing the subject matter experts to create new criteria for competences and abilities that fit the certification scheme better. However, we feel that the benefits of comparability and objectivity outweigh the consequences of the inflexibility. Furthermore, it is beneficial for candidates to work with a single framework and fixed criteria; it makes it easier to show the value of their certificate in an international context. (Or at least, where the e-CF is recognized.)

The Preparation Guide describes all these elements to the candidates and lists the pre-requisites and code of conduct, when applicable. This document, which is freely available, helps comply fully to Article 8 of the ISO/IEC standard.

We have a relative easy job of staying impartial, since we do not train the candidate. Therefore, we comply with the mentioned Article 5.2.1 concerning impartiality. However, being impartial creates a new challenge. We must outsource both part of the creation of the certification scheme to subject matter experts *and* outsource part of the assessment to trainers.

In keeping control whilst outsourcing part of the creation of the certification scheme, the solution is to keep a review by our own experts in the process. This ensures that an EXIN employee ultimately decides on the content of the certification scheme, giving us control.

The responsibility for the assessment part is a little more difficult. Outsourcing the authentic assessment is a threat to keeping full responsibility for the quality of the certification scheme (ISO/IEC standard Article 6.3.2) [2]. We solve this issue by accrediting the training organizations and the trainers. We ensure that the trainers are familiar with best practices for assessing candidates and we regularly inspect the assessments. To ensure that the work of the trainers is in compliance with the exam regulations, we audit the training organizations and we keep records of the audits.

By accrediting training organizations and trainers, we aim to keep high quality assessment and honest assessment. Since trainers are only allowed to change the context of assignments and not the assessment criteria themselves, we keep more control over the assessment, complying with Article 6.3.2. By demanding that the candidate not only shows skills (and is assessed in the training), but also shows their knowledge and insight in a multiple-choice exam, we keep grip on the certification.

The system is not water-tight and we are well aware of that. On the other hand, the system is affordable and easy to implement, even for large-scale assessments in an international context.

## References

1. Mulder, M., Weigel, T., Collins, K.: The concept of competence in the development of vocational education and training in selected EU member states. *J. Vocat. Educ. Train.* **59**, 53–66 (2007). doi:[10.1080/13636820601145549](https://doi.org/10.1080/13636820601145549)
2. NEN-EN-ISO/IEC. NEN-EN-ISO/IEC 17024:2012, IDT - Conformity assessment - general requirements for bodies operating certification of persons (2012)
3. CEN. European e-Competence Framework version 3.0 - a common European Framework for ICT Professionals in all industry sectors, 3rd edn. CEN (European Committee for Standardization) (2014)
4. Borsboom, D., Mellenbergh, G.J., van Heerden, J.: The concept of validity. *Psychol. Rev.* **111**, 1061–1071 (2004). doi:[10.1037/0033-295X.111.4.1061](https://doi.org/10.1037/0033-295X.111.4.1061)
5. Wools, S. Evaluation of validity and validation. In: 9th Annual AEA-Europe Conference on Hisarya, Bulgaria, pp 1–9 (2008)
6. Brennan, R.L. (ed.): *Educational Measurement*, 4th edn. Praeger Publishers, Westport, Connecticut (2006)
7. Govaerts, M.J.B., Schuwirth, L.W.T., van der Vleuten, C.P.M., Muijtjens, A.M.M.: Workplace-based assessment: effects of rater expertise. *Adv. Heal. Sci. Educ.* **16**, 151–165 (2011). doi:[10.1007/s10459-010-9250-7](https://doi.org/10.1007/s10459-010-9250-7)
8. Van Scotter, J.R., Moustafa, K., Burnett, J.R., Michael, P.G.: Influence of prior acquaintance with the rater on rater accuracy and halo. *J. Manage. Dev.* **26**, 790–803 (2007). doi:[10.1108/02621710710777282](https://doi.org/10.1108/02621710710777282)
9. Williams, R.G., Sanfey, H., Chen, X., Dunnington, G.L.: A controlled study to determine measurement conditions necessary for a reliable and valid operative performance assessment: a controlled prospective observational study. *Ann. Surg.* **256**, 177–187 (2012). doi:[10.1097/SLA.0b013e31825b6de4](https://doi.org/10.1097/SLA.0b013e31825b6de4)
10. Engelhard, Jr., G., Myford, C.M., Cline, F. Investigating assessor effects in National Board for Professional Teaching Standards assessments for early childhood/generalist and middle childhood/generalist certification. *Res. Rep. Educ. Test Serv.* I–IV, i-77 (2000)

11. Pitts, J., Coles, C., Thomas, P., Smith, F.: Enhancing reliability in portfolio assessment: discussions between assessors. *Med. Teach.* **24**, 197–201 (2002). doi:[10.1080/01421590220125321](https://doi.org/10.1080/01421590220125321)
12. Clauser, B.B.E., Margolis, M.J.M., Clyman, S.G., Ross, L.P.: Development of automated scoring algorithms for complex performance assessments: a comparison of two approaches. *J. Educ. Meas.* **34**, 141–161 (1997). doi:[10.1111/j.1745-3984.1997.tb00511.x](https://doi.org/10.1111/j.1745-3984.1997.tb00511.x)
13. Van der Schaaf, M.F., Stokking, K.M., Verloop, N.: Cognitive representations in raters' assessment of teacher portfolios. *Stud. Educ. Eval.* **31**, 27–55 (2005). doi:[10.1016/j.stueduc.2005.02.005](https://doi.org/10.1016/j.stueduc.2005.02.005)
14. Ryan, A.M., Daum, D., Bauman, T., Grisez, M., Mattimore, K., Nalodka, T., McCormick, S.: Direct, indirect, and controlled observation and rating accuracy. *J. Appl. Psychol.* **80**, 664–670 (1995). doi:[10.1037/0021-9010.80.6.664](https://doi.org/10.1037/0021-9010.80.6.664)
15. Eva, K.W., Rosenfeld, J., Reiter, H.I., Norman, G.R.: An admissions OSCE: the multiple mini-interview. *Med. Educ.* **38**, 314–326 (2004). doi:[10.1046/j.1365-2923.2004.01776.x](https://doi.org/10.1046/j.1365-2923.2004.01776.x)
16. Iramaneerat, C., Yudkowsky, R.: Rater errors in a clinical skills assessment of medical students. *Eval. Health Prof.* **30**, 266–283 (2007). doi:[10.1177/0163278707304040](https://doi.org/10.1177/0163278707304040)
17. Shechtman, Z.: Agreement between lay participants and professional assessors: support of a group assessment procedure for selection purposes. *J. Pers. Eval. Educ.* **12**, 5–17 (1998). doi:[10.1023/A:1007932515469](https://doi.org/10.1023/A:1007932515469)