Assessing for Situational Judgment
White Paper

Organizations are under increasing scrutiny about how judgment is exercised in the workplace and whether employees will comply with best practices and regulations in the judgments they make. As a result, they are seeing the value of using assessments of situational judgment in the recruitment, onboarding, training and development of their employees, and in strengthening certification programs. In this paper, we discuss how organizations can apply Situational Judgment Assessments, the choices involved in creating them and how Questionmark can support testing

Authors: Eugene Burke
             John Kleeman

Contents

Executive Summary

Is exercising judgment a critical factor in the competence of the employees and contractors who service your organization? The answer is most likely yes. That is why we have developed this paper to introduce you to the opportunities that assessments of situational judgment offer to pre- and post-hire employee programs, and in strengthening certification programs.

We have chosen to refer to these assessments as Situational Judgment Assessments or SJAs[1], as the applications in which we see value from these assessments are broad and extend from recruitment and onboarding to training and certification. We will also discuss how SJAs can be deployed to strengthen assessments and surveys to help reduce regulatory compliance risk.

SJAs offer assessment programs the opportunity to move beyond assessments of what people know (knowledge of what) to assessments of how that knowledge will be applied in the workplace (knowledge of how). This paper does not recommend that organizations and testing professionals adopt SJAs in place of existing knowledge-based assessments. Rather, SJAs offer an opportunity to extend the type of item used to assess for fit to a job or role and in certifying for competence. They can be combined with knowledge tests to assess for both what people know and for what they will do with that knowledge.

Questionmark is well positioned to support SJAs, and we offer this paper as an opportunity for Questionmark clients to explore how they can extract greater value from their investment in Questionmark's assessment management system. For those who are not yet Questionmark clients, this paper offers an opportunity to consider other types of items that can be used to strengthen workplace assessment programs and learn how Questionmark can contribute to effective assessment programs.

The white paper addresses these audiences:

- Psychometricians, testing professionals, work psychologists and consultants who currently create SJAs for workplace use (pre-hire or post-hire) and want to consider using Questionmark technology for such use
- Trainers, recruiters and compliance managers in corporations and government looking to use SJAs to evaluate personnel
- High-tech or similar certification organizations looking to add SJAs to increase the performance realism and validity of their exams

[1] This type of assessment is often referred to as a Situational Judgment Tests or SJT in the employment setting, where SJTs are frequently used in the hiring of new employees. They are also referred to as Situational Judgment Measurements. This white paper adopts the broader term of Situational Judgment Assessment or SJA to reflect the breadth of applications for this form of assessment as explained in the main body of the paper.

1. Why consider assessing for situational judgment

Situational judgment is the ability to draw sound conclusions or make sound predictions about outcomes in a given context. We tend to think of judgment in the workplace as something exercised by those paid to make the big decisions. These are the people who tend to appear in the press headlines when things go wrong. Yet a quick search of occupational databases such as O*Net (https://www.onetonline.org/), using the search words judgment and decision making, instantly throws up 941 occupations ranging from chief executive and secondary science teachers to cooks and cutting and slicing machine setters.

So judgment is not the preserve of the corner-suite executive — it is a competency or skill that permeates all levels of an organization.

This is hardly surprising when you consider that all actions taken in the workplace are the result of someone somewhere exercising judgment. Yet, as a recent joint paper by the Society for Human Resource Management (SHRM) and the Society for Industrial and Organizational Psychology (SIOP) states, the exercise of judgment in organizations is less than perfect, with a failure rate higher than fifty percent[2].

Moving beyond what people know to what people do in assessing for competence

Another reason to consider SJAs is the opportunity for testing and assessment professionals to extend the reach of their assessments beyond the evaluation of what someone knows. Knowing is not the same as doing. SJAs offer a vehicle to assess for higher-level skills and more complex learning objectives, something that has been recognized in the medical profession and that we will discuss later in this paper.

[2] Dalal, R. S., and Bolunmez, B. (2016). Evidence-Based Strategies to Improve Workplace Decisions: Small Steps, Big Effects. Society for Human Resource Management and Society for Industrial and Organizational Psychology. Retrieved via http://www.siop.org/SIOP-SHRM/SHRM-SIOP_Decision_Making.pdf 3 Diagram from Vanderbilt University at https://cft.vanderbilt.edu/guides-sub-pages/blooms-taxonomy/ available under a Creative Commons license. This shows the revised Blooms taxonomy. The original version by Bloom had levels of Knowledge, Comprehension, Application, Analysis, Synthesis and Evaluation.

Consider Bloom's taxonomy, (depicted at right[3]), widely used in learning. The SJAs we will explore in this paper clearly show that this form of assessment offers the opportunity to move beyond the lower levels of this model, "remember" and "understand," and develop assessments that cover the higher levels including "analyze" and "evaluate."

This opportunity is particularly important for certification programs that want to ensure that their assessment of competence is not limited to what people know but extends to assessing what people are likely to do in the workplace — so that their decisions and actions demonstrate appropriate use of knowledge through effective judgment.

To distinguish between knowledge and judgment assessments, consider three item types measuring skills involved in driving a car:

Type 1:   What does a yellow traffic light mean?
    a.   Stop
    b.   Go
    c.   Caution
    d.   Look behind you

Type 2:   If you are driving toward an intersection and the light turns from yellow to red, what should you do?
    a.   Speed up and cross the intersection
    b.   Continue at the same speed and cross the intersection
    c.   Stop suddenly
    d.   Stop gradually

Type 3:   You are giving a close friend a lift to work and because of traffic you are running 15 minutes late for an important appointment. You are driving toward an intersection with traffic lights on it. You can see

the intersecting road to some extent and there seem to be no cars on it. The light turns to yellow as you approach. The car behind you sounds their horn and seems to be accelerating in the expectation that you are going to cross the intersection. Your friend tells you to put your foot down, as you are running late and there are no cars visible on the intersection. What should you do?

a. The road appears to be clear, so speed up and cross the intersection to get through it soon as possible

b. Stay at the same speed and continue to drive through the intersection while keeping an eye out for crossing cars

c. The light is about to turn red, so you should stop immediately and ignore the car behind you as they will need to stop as well

d. Put your hazard lights on to signal to the car behind you and stop at the intersection if the lights facing you turn red

e. Ignore your friend's irritation and move across to the side of the road to let the other car pass

Type 1 is a knowledge question and relies on memory for facts. Type 2 is a question about the application of knowledge and requires some judgment. Type 3 is a question about making a judgment that involves evaluating a situation that has numerous factors (the road layout, traffic light conditions, personal circumstances, other actors involved in the situation) and calls for a choice between more- or less-appropriate responses to that situation.

SJAs are widely used to evaluate how someone would apply softer skills in situations involving others, such as communication, empathy and conflict resolution. Throughout this paper we will explore examples that will show how SJAs can tap into areas of competence that are more difficult to evaluate through knowledge tests alone. Several of the examples are included in a Questionmark sample assessment available at **www.questionmark.com/go/example-sja**.

2. What is an SJA?

An SJA is an assessment or test in which participants are presented with work-relevant situations and asked to make judgments about possible responses. SJAs can be applied in any situation in which a judgment is required to complete a work or other performance task effectively, or where judgment is required to make the most appropriate response to a situation involving others. Those judgments include:

- How an employee interprets and understands a typical work situation in terms of the critical aspects of a situation they should focus on or prioritize
- The correct application of job knowledge in terms of which actions are more versus less likely to achieve a task objective or a desired outcome
- How best to work with and interact with other employees such as members of a work group or team to resolve an issue and overcome obstacles to achieving a shared goal or outcome
- How best to deal with potential dilemmas, particularly those where outcomes are less clear or where information is more ambiguous
- Which responses to typical work situations are more versus less aligned with the expectations of a company, such as compliance with processes and procedures or alignment with company values and organizational norms for behaviour

SJA items differ from other types of items in that they usually include a dilemma. The participant is presented with a choice of responses that usually involve resolving opposing demands in the situation presented.

The structure of an SJA item

Let's take an SJA item developed to assess for a customer-facing role in retail[4]:

You work in a retail store. You have been tasked with refilling a lot of shelves before mid-morning, when the store gets busy with customers. A customer has accidentally broken a bottle of cooking oil. There is broken glass and the oil is quickly spreading across the floor. The customer seems very upset. Other customers are approaching the area where there is broken glass and oil on the floor. One of your colleagues is helping the customer.

What should you do?

a. Watch your colleague for a moment to see if your help is needed
b. Help your colleague deal with the situation by calling the cleaning staff
c. Stay focused on completing your current task of re-filling the shelves in time

The dilemma in this item is that while you, the participant, have been tasked with restocking the shelves before a busy shopping time, there is a situation unfolding with the broken bottle of oil that could escalate and impact other customers.

The response choices are effectively to do nothing yet (A), get involved at the risk of your main task being delayed (B) or continue with your main task and ignore the situation (C). Each of these choices has a consequence. The judgment here is which of these choices is likely to lead to a preferred or better outcome and which is likely to lead to a less preferred or a worse outcome.

For this item, A is a neutral option in that it is unlikely to improve the situation or make it worse. Option C relates to the primary task given to the participant at the beginning of the situation description (the item stem), but choosing this option could make the situation worse if the incident with the broken bottle of oil and the upset customer were to escalate.

[4] Taken from Burke, E., Vaughan, C., and Fix, C. (2009). Situational Judgement Test (SJT) Workshop. Workshop delivered at the European Congress of Psychology, Oslo, July.

Option B means that the participant would support their colleague in dealing with the broken bottle of oil and the upset customer. This option also gives the participant the opportunity to continue with their main task of restocking the shelves later.

The figure below summarizes the structure of an SJA item:
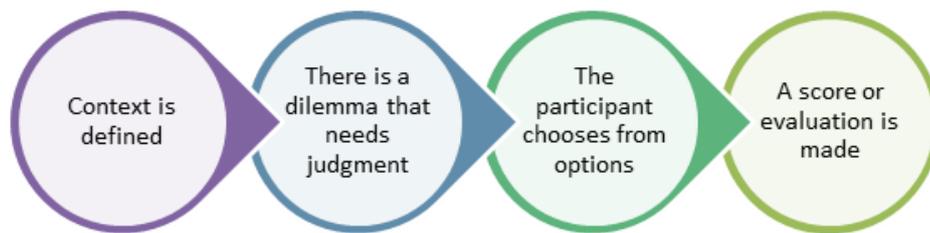
Let's explore this structure in another situation, involving a technical team[5]:

You work as part of a technical support team that produces work internally for an organization. You have noticed that often work is not performed correctly or a step has been omitted from a procedure. You are aware that some individuals are more at fault than others as they do not make the effort to produce high quality results and they work in a disorganized way. What do you see as the most effective and the least effective responses to this situation?

    a. Explain to your team why these procedures are important and what the consequences are of not performing these correctly.
    b. Try to arrange for your team to observe another team in the organisation who produce high quality work.
    c. Check your own work and that of everyone else in the team to make sure any errors are found.
    d. Suggest that the team tries many different ways to approach their work to see if they can find a method where fewer mistakes are made.

[5] Taken from Burke, E., Ablitt, H., Guillard, G., and Vaughan, C. (2012). Global, secure, valid and relevant: Four breakthroughs in the science and practice of SJTs. Symposium delivered at the British Psychological Society's Division of Occupational Psychology Annual Conference, Chester, January.

This SJA item is more complex as the participant now has four options to choose from and is being asked to distinguish among the most and least effective responses. This SJA item format is often used to assess for finer discrimination in the participant's analysis of the situation.

Option C is a way to deal with the errors but places a significant amount of work on the participant.

It also does not address the behavior of the other members of the team. Option B seems reasonable but will not address the issues immediately (access to another team must be arranged) and may not address the team's disorganized approach. These two options are unlikely to make the situation worse, but they are not guaranteed to address the team's behavior.

Option D is asking a disorganized team to engage in an open and undefined set of experiments that could increase rather than reduce errors in the work produced. This is, on balance, the least effective of the options presented.

Option A does require some confidence in dealing with the potential pushback from the other team members. However, it does address the team's behavior directly and it does approach the problem in terms of consequences for the team as a whole rather than as a subjective criticism of the team and any individual. This, therefore, is the option that is most likely to have a positive effect.

Assessing for high-stake situations

SJAs have been developed for highly technical roles to capture high-stakes situations that may confront people in those roles. For example, the SJA below[6] presents a dilemma to an aviation pilot who must make a judgment about the most effective option for responding to the situation presented:

[6] Hunter, D. R. (2003). Measuring general aviation pilot judgment using a situational judgment technique. International Journal of Aviation Psychology, 13, 373–386.

> You are flying an "angel flight" with a nurse and noncritical child patient to meet an ambulance at a downtown regional airport. You filed visual flight rule (VFR). It is 11:00 p.m. on a clear night when, at 60 nautical miles (nm) out, you notice the ammeter indicating a battery discharge and correctly deduce the alternator has failed. Your best guess is that you have from 15 to 30 minutes of battery power remaining. You decide to:
>
> a. Declare an emergency, turn off all electrical systems, except for 1 NAVCOM and transponder, and continue to the regional airport as planned
> b. Declare an emergency and divert to the Planter's County Airport which is clearly visible at 2 o'clock at 7 nm
> c. Declare an emergency, turn off all electrical systems, except for 1 NAVCOM, instrument panel lights, intercom and transponder, and divert to the Southside Business Airport which is 40 nm straight ahead
> d. Declare an emergency, turn off all electrical systems, except for 1 NAVCOM, instrument panel lights, intercom and transponder, and divert to Draper Air Force Base which is at 10 o'clock, at 32 nm

In case you are not a pilot and are wondering what the most effective answer is, it is B. This is the nearest airport, is easily visible to the pilot (2 o'clock is slightly to the right of the pilot's line of direct sight) and requires minimal course correction to reach. While understanding this item may require some knowledge of aviation, it shows how SJAs can be developed for high-stakes situations in highly technical roles[7].

This feature of SJAs has been recognized in the processes used to select participants for the Foundation Programme in the UK, a two-year post-graduate medical training programme that forms the bridge between medical school and specialist/general practice training[8]. SJAs are used because professionalism and interpersonal skills are central to the role of a doctor.

---

[7] This form of SJA has continued to be developed as a tool for assessing the judgment of general aviation pilots.
[8] Patterson, F., Ashworth, V., and Good, D. (2013). Situational Judgement Tests: A Guide for Applicants to the UK Foundation Programme. Medical Schools Council. Retrieved via
http://www.foundationprogramme.nhs.uk/download.asp?file=Situational_Judgement_Tests_Monograph_FIN AL_August_2013-1.pdf

Following a detailed job analysis, five attributes of professionalism were identified as the basis for constructing the SJTs now used to screen applicants to the UK Foundation Programme:

Attribute	Examples
Commitment to Professionalism	• Dealing with issues of confidentiality such as hearing a colleague talking about a patient outside of work • Challenging inappropriate behavior such as a consultant speaking to a colleague/patient in an inappropriate way • Commitment to learning such as the need to go to teaching while also being needed on the ward
Coping with Pressure	• Knowing how to respond when you make a mistake such as providing wrong medication to patient • Dealing with confrontation such as an angry relative • Seeking help when not sure of the correct procedure/best way of doing things
Effective Communication	• Gathering information and communicating your intentions to nursing staff or other colleagues • Negotiating such as for a scan from radiology • Listening and effectively communicating such as with an unhappy patient or relative
Patient Focus	• Identifying that a patient's views and concerns are important and they should have input into their care • Considering that a patient may have different needs from others around them • Spending time trying to understand a patient's concerns and empathising with them

Working Effectively as Part of a Team	• Recognising and valuing the skills and knowledge of nursing staff, when faced with a disagreement about a patient's care • Consulting with colleagues about how to share workload fairly • Offering assistance and support to colleagues when they are not able to handle their workload

We have spelled out these attributes in full as an illustration of a framework from which SJA dilemmas could be identified. Individual fields of work will have different attributes and examples but there will often be similarities. For example, an accountant may need to flag an issue with a more senior colleague if he or she is concerned about a potential audit risk, or an IT consultant may need to deal with the pressures and communication challenges around systems changes.

Here is an example of an SJA item developed using this framework:

You are looking after Mr. Kucera who has previously been treated for prostate carcinoma. Preliminary investigations are strongly suggestive of a recurrence. As you finish taking blood from a neighboring patient, Mr. Kucera leans across and says, "Tell me honestly, is my cancer back?"

Rank the appropriateness of the following actions in response to this situation:

a. Explain to Mr. Kucera that it is likely that his cancer has come back
b. Reassure Mr. Kucera that he will be fine
c. Explain to Mr. Kucera that you do not have all the test results, but you will speak to him as soon as you do
d. Inform Mr. Kucera that you will chase up the results of his tests and ask one of your senior colleagues to discuss them with him
e. Invite Mr. Kucera to join you and a senior nurse in a quiet room, and get a colleague to hold your 'bleep' and then explore his fears

The correct ranking for this item is DCEAB. Responses A and B would be founded on incomplete information as the investigations are preliminary and are only

suggestive of the patient's condition. They may be misleading, and A will not deal with any anxiety the patient is experiencing. E may address the patient's anxieties but leaves the diagnosis of the patient's condition inconclusive. C is a more honest response but begs the question of whether the doctor responding is qualified to provide a diagnosis. D commits the doctor to resolving the situation and to ensuring an appropriate level of expertise is involved in providing a diagnosis.

Participants like the SJA format and it works

A wide body of research has shown that SJAs are well accepted and received positively by participants[9], are fairer than many other forms of assessment[10] and predict real-world workplace outcomes and job performance[11]. Simply put, the way an individual responds to SJAs has been shown to predict actual behavior and performance in the workplace.

In the next section of this paper, we will explore various applications of SJAs in more detail before we explain how to develop effective SJA items.

[9] Lievens F, Sackett PR. Situational judgement tests in high-stakes settings: Issues and strategies with generating alternate forms. Journal of Applied Psychology, 2007, 92, 1043-1055. See also Ployhart, R. E., and Ward, A. (2013). Situational Judgment Measures. For a discussion of deploying SJAs in multinational contexts, see Burke, E., Vaughan, C., and Glennon R. (2013). Employment testing and assessment in multinational organizations. Both chapters are in K. F. Geisinger (Ed.). APA Handbook of Testing and Assessment in Psychology. Washington DC: American Psychological Association.

[10] Whetzel, D. L., and McDaniel, M. A. (2009). Situational judgment tests: An overview of current research. Human Resource Management Review, 19, 188–202.

[11] Chan, D., and Schmitt, N. Situational judgment and job performance. Human Performance, 15, 233-254.

3. Applications of SJAs

In this section, we will explore how SJAs can be applied by organizations in their employee programs, pre- and post-hire, how SJAs are being deployed in certification programs, and how SJAs can be applied by compliance professionals to better understand the impact of their programs and to identify where further improvement is required.

Pre-hire and helping employers and job applicants make better decisions

As recruiting new employees has become more competitive and improving the return from investments in hiring processes has increased in importance, many organizations in the private and public sectors have turned to a form of SJA referred to as a Realistic Job Preview or RJP.

RJPs are deployed early in the recruitment process to give a prospective employee a realistic view of what a job involves as well as the opportunity to try the employing organization on for size. This is essentially about helping the job applicant determine whether the organization is likely to be a good fit for them (sometimes referred to as fit-to-place) and whether the job or role is a match for their interests and aspirations (sometimes referred to as fit-to-job or fit-to-role).

SJAs offer an approach to RJPs that allow job applicants to understand what is expected of new employees and the types of situations and challenges they might experience in their jobs.

Below is an example of an SJA developed as an RJP for a graduate recruitment program. While we will share this example in the form of text, it could be presented in a graphical form using images, cartoons or video. There are choices in using these media such as expense and the transportability of content if the RJP is going to be used in different countries and in different languages.

The job applicant lands on the recruitment site and is directed to five or six SJA items that present situations such as the following:

> You are working as part of a project team on a client site. It's the start of the day and you have arrived early on site. So far, the project has gone well and is on time, but your team leader has just called to say they will not make it today due to illness. You are approached by a senior client who has asked for a meeting in thirty minutes to be briefed on the progress of the project. You are one of several team members but the other team members have not arrived yet.
>
> What should you do?
>
> a. Agree to the client's request and do your best to prepare a brief for the client
> b. Tell the client that your colleagues will be arriving shortly and it would be best to discuss this with them so that the client will get a full briefing
> c. Tell the client that your team leader is off sick and that this will have to wait until the team leader is able to come on site

This is a lower-stakes application of SJAs. The purpose of the RJP in this instance is to provide feedback on how the organization would like its employees to handle situations such as this. For example, here is what the feedback on each option might be:

A. "We appreciate employees who show initiative but we are also very focused on ensuring that our clients get the best and most complete information they need."

B. "That looks like a sensible way to go even if sometimes our clients can be very insistent on quick responses. The other members of the team will have information the client will need and may have more experience in dealing with situations like this."

C. "That may not satisfy the client and it might be best to share your team leader's situation with the other members of the team when they get there so that the team can then decide how best to proceed."

Some organizations allow applicants full view of the feedback on each option but ask them to still choose the option that describes how they would typically react to a situation like this. Five or six SJA items as shown above can be sufficient to give the applicant an insight about whether the job or organization is likely to be a good fit for them. If all SJA items used to provide RJPs are scored, then the overall feedback might range between "We look like a good fit for you. Click here to proceed to the next stage of the application process" to "You might want to think about whether we would be a good fit for you. Click here if you would like to proceed to the next stage of the application process."

Research on RJPs has shown that well-constructed RJPs help promote job satisfaction and intention-to-stay among new hires[12]. While SJAs in the form of RJPs have been typically used in pre-hire, there are post-hire applications for this use of SJAs as organizations look to develop in-house recruitment processes offering existing employees a more engaging way to explore other opportunities with their current employer.

Scored SJAs will often appear in later stages of the hiring process to inform final hiring decisions, as illustrated by the examples of items for retail employees, technical teams and medical students we shared earlier in this paper. Now let's explore post-hire applications of SJAs.

---

[12] See Earnest, D. R. (2015). Realistic Job Previews. In Wiley Encyclopedia of Management. Volume 5: Human Resource Management. John Wiley and Sons

Post-hire and using SJAs in workforce training and development

While many would acknowledge that the transfer of learning into the workplace is enhanced by the opportunity to practice the application of knowledge and skills in realistic situations, the challenge with providing those opportunities is the simple factor of cost. This is particularly true of scenariobased training, simulations or training that relies on role playing[13].

Many organizations use observational assessments (e.g., using Questionmark technology) to help assess work skills including those that use judgment, but these also need an observer to rate the participant.

Another advantage of SJAs is that they offer a relatively low-cost and scalable way to deliver ondemand opportunities for trainees and learners to try out their understanding against short vignettes, obtain feedback and develop competence[14]. Again, this is not an "either or" recommendation, as SJAs deployed at suitable points in a training program can be used to supplement and complement other training.

Here are three ways in which SJAs can add value to training and development programs:

---

[13] For a discussion of how SJAs can add value to scenario based training programs, see Fritzsche, B. A., Stagl, K. C., Salas, E., and Burke, C. (2006). Enhancing the design, delivery, and evaluation of scenario-based training. Can situational judgment tests contribute? In J. A. Weekley and R. E. Ployhart (Eds.) Situational Judgment Tests: Theory, Measurement, and Application. Mahwah, New Jersey: Lawrence Erlbaum Associates

[14] For a discussion of how SJAs can provide a low-cost, reliable and agile way to meet training evaluation needs see Gunter, S., Mellone, I., Oakley, K, and Faben, C. (2013). Evaluate Training and Performance Effectively, Quickly, and Inexpensively Using the Situational Judgment Test (SJT). Interservice/Industry Training, Simulation and Education Conference (I/ITSEC). Orlando, Florida, December.

- **Diagnosing the learner's development and training needs.** Much as with the RJPs mentioned in pre-hire applications, SJAs offer a means to engage learners in a variety of situations that typify the need for specific knowledge and skills or, at more advanced stages of a program, a mix of knowledge and skills as the learner becomes comfortable with more complex and demanding situations. SJA items can be scored to provide targeted feedback about where the learner needs to invest their efforts and to guide them to suitable training resources in support of achieving their learning goals. Used as pre-training diagnostics, SJAs can be used to pinpoint skills and development gaps to guide both learners and trainers in leveraging training investments more effectively

- **Delivering training experiences with feedback.** SJAs can also be deployed as training materials in themselves and, given their comparative ease and lower cost of authoring and deployment, offer the opportunity to build breadth and variety into the learner's training resources. They also offer the opportunity to share and develop skills that rely on tacit rather than explicit knowledge such as the softer skills essential to roles in customer service and sales, including effective communication, negotiation, problem solving for a customer and dealing with complaints. With organizations investing heavily in diversity and inclusion programs, we see a strong value in the deployment of SJAs to support those programs and, as we will describe below when we consider the use of SJAs in surveys, to evaluate whether those programs are communicating and driving desired attitudes and behaviors.

- **Evaluating training effectiveness.** SJAs are also used to test for whether learner outcomes have been achieved at both the level of the individual learner and for cohorts of learners where the focus is on the effectiveness of the training program as a whole. We do not advocate that SJAs be used as the sole method of assessment but rather that they be used with other methods of assessment to ensure that all key aspects of learning are evaluated. However, for programs that rely heavily on assessment of knowledge alone, SJAs offer a low-cost entry point to improving the validity and robustness of training and development evaluations.

Training needs frequently change as new products, services, technologies and procedures are developed and rolled out within organisations and to their clients. With their lower cost of authoring and deployment, SJAs offer a more agile and efficient means of developing and adapting training and development resources to match an organization's changing needs.

What might an SJA used in this context look like? Here is an example adapted from law enforcement, where SJAs are used in a program for developing high-potential senior police officers. The item relates to community and press relations:

You are a senior officer giving a live interview to the press about the effect of recent budget cuts on policing. The City Council Supervisor and Chair of the Community Safety Partnership has recently taken a substantial pay increase whilst at the same time announcing redundancies (layoffs). This has not gone down well with employees or the local media. The Supervisor is a highly effective and well-regarded colleague and the Community Safety Partnership has achieved good results. The interviewer asks you for your views on the matter. Please rate the effectiveness of the following options in responding to the situation outlined above.

a. Tell the interviewer that you do not think that senior public servants should accept pay increases in these difficult times, but that you do not have all the facts on which to judge this case

b. Politely inform the interviewer that the issue of the Supervisor's pay is a matter for him and his employer and that you do not wish to comment

c. Say that you understand the difficult decisions that are being taken in all public sector organisations and that the Supervisor's acceptance of the pay rise sends out a negative message in these difficult times

d. Say that the Supervisor is an effective performer but that you understand there has been significant ill feeling in the community over this matter and you are sure that the relevant authority will want to give an explanation

e. Tell the interviewer that you are here to speak about police matters so you do not wish to comment on this issue and tell them they will need to speak to the relevant authority

How does this SJA item differ from a scenario-based question? Scenario-based questions tend to be designed to ensure that a participant understands the correct sequence of actions required to execute a task or a process. In the example above, although there is likely to be a protocol for a senior officer to follow in liaising with the press, the SJA item explores how a participant would navigate circumstances that could lead to a compromise of that protocol and escalate the situation.

SJAs in certification programs

Competence can be defined as the combination of training, skills, experience and knowledge that a person has and their ability to apply them to perform a task or achieve a work goal effectively. We have already raised the challenges of assessing for the application of experience and knowledge, which is an expensive enterprise when observational or interview methods are used — methods that themselves bring challenges in avoiding subjectivity and bias.

Let's consider a hypothetical example. ABC Corporation is a software vendor. ABC Corporation develops a certification program for its own engineers, its partners and its customers to encourage good practice. ABC encourages customers to use certified people to manage and deploy the application. ABC's certification program uses questions on how the software works and on good practices. These include drag-and-drop questions that require participants to mimic tasks within the software. ABC's certification program is well respected.

Over time, feedback on the software shows that customers get more value out of the application, are more satisfied and are more likely to renew and increase their spend on the company's products and services if the software is deployed, configured and managed effectively. Comments from the vendor's engineers, partners and customers suggest that, although the knowledge exams that are the basis for the current certification are useful in ensuring that those certified have a good understanding of the software, there is more to deploying ABC's software effectively than just knowing how it works.

This feedback points to the fact that judgment is needed to be able to configure the software to meet individual customer needs. The feedback also highlights that significant factors in customer satisfaction are how engineers work with customers and how they deal with those situations when the aspirations of customers cannot be met by the software. The outcomes of those interactions are significant in building customer loyalty, satisfaction and referrals.

ABC sets up a review with internal stakeholders (engineers, customer service agents and sales people) to identify a shortlist of the factors that distinguish more successful from less successful customer engagements. They identify five themes: getting to know the customer, framing the customer interaction, meeting the customer's needs, dealing with out-of-scope requests and anticipating future customer needs.

To build these aspects of competence into the certification program, ABC decides to include SJAs within certification exams alongside knowledge items, creating a two-part exam. Feedback from key stakeholders shows that the new SJA items are seen as adding to the fidelity of the exam by broadening its relevance to real-world job skills.

While that is a hypothetical example, it reflects real-world concerns that have emerged in the hightech certification field as well as in the medical profession, law enforcement and protective services — with many accepting that certification based on knowledge alone is inadequate. Let's follow those concerns through via an example of an SJA item designed to assess for effective judgment of IT security professionals[15]:

[15] The Questionmark security team advise that they consider option A to be the least effective approach as it fails to look at how the hacker got into the system and leaves a door open. They suggest option C is the most effective option. Option B is possible but does not seem to take the situation seriously enough (given a hacker has got access to a payment processing application), and option D may not resolve the situation quickly enough for the organization's benefit.

You are responsible for the security of your company's payment processing application. You have recently received an email from a person who claims to have hacked into the system and has discovered that you are running an out-of-date and vulnerable operating system on one of your servers. You have checked and the hacker is correct. Your preliminary check has shown no trace of how or whether the hacker got into your system. Those checks suggest that there isn't anything is unusual with the systems performance.

From the options below, identify the actions that would be most and least appropriate in dealing with this situation.

  a. Be grateful that this is a white hat hacker who has done the company a favor by identifying the vulnerable operating system. Your main action should be to fix the operating system as soon as possible and ensure that operating system reviews and updates are done more quickly
  b. Drop your other work and run a thorough system scan and log review to identify if anything else is awry with the system. If nothing is found, update the operating system and continue normally
  c. Shut down the system and call in security consultants to do a thorough review of what intrusion there has been and whether there has been any data breach, financial loss or other serious impact, and then update the operating system
  d. Fix the operating system and then report the security breach and the hacker's email to the police so that any potential criminal activity can be acted on by the relevant authorities

Much as in our hypothetical case of ABC Corporation, including an SJA item such as this in a certification exam serves not only to test whether someone certified is likely to demonstrate sound judgment in the workplace, and therefore competence, but also serves to communicate the expectations for demonstrating best practice.

From this perspective, SJA items offer a means to communicate that certifying organizations are abreast of current issues and acting to promote best practice as well as addressing real-world risks and threats to organizations and their customers.

SJAs in support of compliance programs

The challenge to those charged with compliance in organizations was recently captured in an EY paper[16]:

> *Rules are immensely important in defining the parameters within which teams and individuals should operate.* ***But rules alone are not enough. It's whether and how they are adopted by people when making decisions that matters****. Even the most well-intentioned rules can be overwhelmed by culture when it is ignored. The key is striking the right balance to create a space where rules and culture reinforce one another. To foster the cultures that can effectively nudge people to engage in desired behaviors, organisations have to start by understanding the cultures that they currently have. It's hard, but critical work if organisations are to gain leverage over how their people exercise judgement.* ***And judgement is at the heart of nearly every business scandal that ever occurred[17]****.*

As this quote shows, regulatory authorities in banking, finance and other industries are fast coming to the realization that the processes, procedures and audits that many organizations have in place are unlikely to safeguard against noncompliance, and the discussion around improving defense against noncompliance is moving rapidly to what organizations know about the behaviors of their employees.

For legal reasons as well as good business practice, many organizations survey their employees to track risks of noncompliance. The weakness in current practice for these surveys is that many fail to adopt content that provides insight on the judgment of employees and how they are likely to act when confronted with situations that challenge compliance and may threaten the legal and financial integrity of the organization.

---

[16] Ernst & Young (2016). The Route to Risk Reduction: Better Rules of Better Decisions? Retrieved via
http://www.ey.com/Publication/vwLUAssets/EY-the-route-to-risk-reduction-better-rules-or-betterdecisions/$FILE/EY-the-route-to-risk-reduction-better-rules-or-better-decisions.pdf
[17] The emphasis in bold is by the authors and not the original paper.

Let's explore two SJA items that would be suitable for embedding in an organizational survey or as standalone items for a short pulse survey. SJA items such as this can be deployed to evaluate a range of organizational programs including safety and diversity and inclusion. Our first example is set in the context of customer approvals. For this item, we ask the respondent to indicate what they see as a more or less effective way of addressing the situation presented. This is a "should" question directed at whether the respondent would demonstrate appropriate judgment and therefore meet the expectations for good performance.

Our second example is framed more as a "would" question as it is asking the respondent to indicate how they figure employees in the organization would typically act. We will come back to the distinction between "would" versus "should" in SJAs in the next section of this paper when we discuss how to develop SJAs, but we are mentioning that distinction here to show the flexibility that SJAs offer when deployed as survey content.

For the purpose of a survey, and particularly when survey responses are not attributable to a specific employee (i.e., the responses are submitted anonymously), SJA items can be deployed to gather data akin to a poll with results showing the percentage of respondents who choose specific options or, as in our second example, rate different options in terms of how frequently respondents believe that employees in the organization would act.

Data of this kind segmented by attributes of respondents (such as business function, location and seniority as well as demographics such as age, gender and tenure) serve to indicate whether compliance will be acted out in the workplace, evaluate where compliance professionals need to focus their efforts and evaluate whether compliance programs are gaining traction. Given that more detailed actions would normally follow from the insights provided by such SJA data, the use of single SJA items in this context is justifiable and defensible. This contrasts with using SJAs in tests or exams, where the use of SJAs likely would need to meet the requirements of coherent measurement scales, combining data across items and meeting psychometric standards expected of

assessments that make higher-stakes decisions about individuals. Let's explore our two survey examples.

SJA as Survey Content: Example 1

You work in the back office in the team approving new customers, ensuring that the organization's procedures have been followed (such as credit rating and know your customer). Your manager is away on holiday this week. A senior manager in the company (three levels above you) comes into your office and says that there is an important new customer who needs to be approved today. They want to place a big order, and he can vouch that the customer is good. You review the customer details, and one piece of information required by your procedures is not present. You tell the senior manager and he says not to worry, he is vouching for the customer. You know this senior manager by reputation and have heard that he got a colleague fired a few months ago when she didn't do what he asked.

You would:

a. Take the senior manager's word and approve the customer
b. Call your manager's cellphone and interrupt her holiday to get advice
c. Tell the manager you cannot approve the customer without the information needed
d. Ask the manager for signed written instructions to override standard procedures to allow you to approve the customer

SJA as Survey Content: Example 2

Bill has just joined a new sales team and has completed his first business trip. Bill submits his expenses claim to Jane, his manager, and Jane comments that the claim must be wrong as it underestimates typical expenses for that trip. Jane insists that Bill re-submit the claim and add 25% to bring it into line with what Jane was expecting. Rate how you think people in the organization are likely to behave in a situation like this. Use the following scale to rate each of the options below (enter one of the following numbers against each option):

1 = Very Unlikely 2 = Unlikely 3 = Neutral 4 = Likely 5 = Very Likely

a. Bill toes the line because Jane has more experience about what this expense claim should be and she knows the company's processes better
b. Bill goes back and double-checks his expense claim and, if his original claim was correct, resubmits it as is and tells Jane it is accurate
c. Bill asks other colleagues what they think because maybe this is accepted practice and Bill should not go against it
d. Bill contacts the finance department and checks what the expenses for this trip typically are before resubmitting the claim to Jane

Now that we have explored a range of applications of SJAs, we will discuss how to develop effective SJAs.

4. Constructing SJAs

Developing effective SJA content is twofold: it involves steps specific to the creation of typical workplace situations and meaningful response options as well as steps common to the development of all effective tests, exams and assessments.

Higher-stakes versus lower-stakes SJA applications

The applications of SJAs we explored in the previous section of this paper can be broadly categorized into lower- and higher-stakes scenarios. Lower-stakes applications refer to scenarios in which SJA content is used to help inform someone — such as with RJP or employee onboarding applications — or to gather data to help inform policy and guide actions — such as in deploying SJA content for compliance surveys. Higher-stakes applications include the use of SJAs to make decisions on whether to make a job offer or to declare someone as competent to undertake workplace tasks — such as successful completion of a training program or a qualification obtained through a certification program.

Whether lower-stakes or higher-stakes, the quality of SJA content will depend on efforts to construct effective SJA item stems and response options as well as the other key steps in developing SJA content that will we describe next. For higher-stakes applications where SJA items will be developed to report scores to be used in hiring or competence decisions, additional steps should be followed to ensure that SJA items and scales have adequate properties and meet best practice standards as set by professional bodies[18].

---

[18] For example, the Standards for Educational and Psychological Testing from AERA, APA and NCME.

In many applications, SJA items will be combined into part or all of an assessment. Data across SJA items are aggregated to obtain SJA scores, which may be total scores (sum of scores across all items) or sub scores (the sum of scores on subsets of items, for instance a topic). In either case, we advise that SJA programs follow the steps recommended by best practice guidelines for gathering evidence that scores are reliable, accurate and valid[19].

Key steps and considerations in developing SJA content

Let's explore the more specific steps involved in developing effective SJA content. To illustrate these steps, we will use an SJA item, shared earlier in this paper, that was used to screen for a medical doctor program[20]. We will deconstruct this item to show how SJA items are built up and how different response formats and item scoring methods can be applied to an SJA item.

---

[19] In common with other types of tests and exams, item parameters to be examined from pilots and preoperational trials of SJA content will include item difficulties (percentage of candidates selecting the correct response option), distractor analysis (proportions responding to alternative incorrect responses), item-total correlation discrimination or similar, checks on scoring keys as well as scale (total score) reliabilities and total and/or partial score distributions. Evidence for higher-stakes applications will be stronger when that evidence includes correlations with outcomes such as job performance and when items and scales have been analyzed for potential bias by demographic such as age, gender and ethnicity

[20] Note that very similar items have also been developed to screen applicants for the position of hospital pharmacist, which again shows how SJA content is gaining traction across a range of medical specialties.

Here is that item:

You are looking after Mr. Kucera who has previously been treated for prostate carcinoma. Preliminary investigations are strongly suggestive of a recurrence. As you finish taking blood from a neighboring patient, Mr. Kucera leans across and says, "Tell me honestly, is my cancer back?"

Rank the appropriateness of the following actions in response to this situation:

  a. Explain to Mr. Kucera that it is likely that his cancer has come back
  b. Reassure Mr. Kucera that he will be fine
  c. Explain to Mr. Kucera that you do not have all the test results, but you will speak to him as soon as you do
  d. Inform Mr. Kucera that you will chase up the results of his tests and ask one of your senior colleagues to discuss them with him
  e. Invite Mr. Kucera to join you and a senior nurse in a quiet room, and get a colleague to hold your 'bleep' and then explore his fears

- **Developing SJA Item Stems**

The first step in constructing SJA content is to conduct an analysis to identify situations that reflect the demand for a given domain of knowledge, skill, and work behavior, competency or compliance action. For most of the applications we explored in the previous section of this paper, this will require a job analysis. Whether conducted as a fresh exercise or in the form of existing materials, decisions need to be made as to which dimensions identified will be used to frame SJA items. Existing materials may not list in detail situations that typify how skills, knowledge or competencies will be applied in the workplace, so additional steps may be required involving Subject Matter Experts (SMEs) to obtain examples of situations that demand the use of the target skills, knowledge or competencies that the SJA will be used to assess for

Two approaches to job analysis are typically used to create SJA content:

- **Critical Incidents Analysis** can be executed through online surveys as well as workshops with SMEs to identify situations that discriminate between effective or appropriate work behaviors and skills (observable actions) and less effective or appropriate work behaviors and skills. The situations generated are then analyzed to identify core themes and to construct a limited set of dimensions to frame the development of SJA content. The key here is to identify those situations that describe the application of critical work behaviors and skills, and how those behaviors and skills demonstrate competent or less competent responses to situational demands. Once a defined set of dimensions has been obtained, situational content is then sorted by dimensions to provide the bare bones of the stem and response options that will become SJA items.

- **Competency-based Approaches** can be applied when competencies for a job or role have already been defined. Again, surveys and/or workshops may still be required to elicit situations in which SMEs have observed either effective/appropriate versus ineffective/inappropriate responses to situations demanding the target competency. SMEs are presented with a definition of a target competency and asked to provide an example of when they have seen someone demonstrate that competency effectively. In a workshop setting, SMEs can share the situations generated to discuss which of them best reflect the demand for the target competency. This helps the facilitator develop a shortlist of situations most relevant to the target competency. SMEs are then asked for examples of the observed less-effective demonstration of the competency or situations when the competency was demanded but someone failed to demonstrate it. In addition to generating the basis for SJA content, cycling through competencies and situations in this way also serves to build understanding among SMEs of the SJA program's objectives and strengthen their engagement in the process. This approach works best when it is focused on five or six competencies to avoid SME fatigue and safeguard the quality of the content generated.

Let's return to our example item. Whether generated using critical incidents or a competency-based approach, the bare bones of the situation is a patient who is

concerned about whether their condition has returned. Earlier, we shared the target attributes of this and other SJA items developed for the medical doctor screening program.

Let's assume that this situation is generated to reflect the demand for the following attribute or competency[21]:

Attribute	Examples
Patient Focus	• Identifying that a patient's views and concerns are important and they should have input into their care • Considering that a patient may have different needs from others around them • Spending time trying to understand a patient's concerns and empathising with them

- **Constructing a Dilemma for the Item**

Stem At this point, we have a target attribute and the bare bones of a situation. The next step in developing effective SJA content is to evolve the bare bones of the situation to include a dilemma in the form of situational demands and choices that will be presented to the participant in the form of response options. Focusing on the item stem for now, we can see from the target attribute that the situational demands should emphasize how patients' concerns are recognized and acknowledged — and how empathy is shown while ensuring that the specific needs of a patient are met. Let's return to the item stem from our example to see how the bare bones of the situation have been evolved to present situational demands related to the target attribute and to set up the choices the participant will be asked to make when they record their answers to the response options (which we will come to presently):

---

[21] The assignment of this situation to this attribute is offered solely for the purposes of illustration and may not reflect the original analysis conducted by the authors of the UK Foundation Programme SJA. The authors of this paper saw the assignment to this attribute as a reasonable inference.

You are looking after Mr. Kucera who has previously been treated for prostate carcinoma. Preliminary investigations are strongly suggestive of a recurrence. As you finish taking blood from a neighboring patient, Mr. Kucera leans across and says "Tell me honestly, is my cancer back?"

We can infer from the patient's question that he is likely to be anxious about whether his condition has returned. The possibility of this is mentioned earlier in the stem as is the patient's history. The situation is placed in an area where other patients are present (the "neighboring patient"). While more complex SJA stems can involve a variety of conflicting demands and more ambiguous information, this example shows that an effective SJA stem must act to trigger or elicit a judgment related to a specified skill, competency or compliance challenge.

- **Developing Response Options**

  Let's assume that we have collected information from workshops with SMEs and we have a list of what would be more and less effective or appropriate behaviors or actions against the target attribute. We can now explore examples of those behaviors and actions to shortlist those most suitable for the situation described in the item stem, and contextualize our shortlist of potential responses so that they tie directly to the item stem.

  Once we have drafted a number of SJA items, it is useful to test the content with a small group of experts with relevant domain knowledge. For lower-stakes applications, this check may be sufficient prior to deploying content. In higher-stakes applications, this step would normally precede piloting and statistical checks. It would identify problematic SJA items early and ensure that the investment in piloting is maximized, to check proposed scoring keys function as expected, and on the degree of consensus among SMEs to ensure that answer keys do reflect standards expected for the demonstration of workplace competence.

This content check can take the form of a short survey that presents SMEs or content experts with the item stem and proposed response options, and asks them to rate the response options by effectiveness or appropriateness.

Here is how our example item might appear in such a content review exercise:

> You are looking after Mr. Kucera who has previously been treated for prostate carcinoma. Preliminary investigations are strongly suggestive of a recurrence. As you finish taking blood from a neighboring patient, Mr. Kucera leans across and says "Tell me honestly, is my cancer back?"

Target Attribute/ Skill/ Competency	**Patient Focus** • Identify a patient's views and concerns are important and are input into their care • Consider that a patient may have different needs from others around them • Spend time to understand a patient's concerns and empathise with them
Situation	You are looking after Mr. Kucera who has previously been treated for prostate carcinoma. Preliminary investigations are strongly suggestive of a recurrence. As you finish taking blood from a neighboring patient, Mr. Kucera leans across and says, "Tell me honestly, is my cancer back?"

Rate the options shown below in terms of how effective they are in responding to this situation where 1 = Very Ineffective, 2 = Ineffective, 3 = Neither Ineffective/Effective, 4 = Effective and 5 = Very effective

Option	Description	Your Rating
A	Explain to Mr. Kucera that it is likely that his cancer has come back	
B	Reassure Mr. Kucera that he will be fine	

C	Explain to Mr. Kucera that you do not have all the test results but you will speak to him as soon as you do
D	Inform Mr. Kucera that you will chase up the results of his tests and ask one of your senior colleagues to discuss them with him
E	Invite Mr. Kucera to join you and a senior nurse in a quiet room, and get a colleague to hold your 'bleep' and then explore his fears

- **"Should" versus "Would"**

Instructions Before we move onto how an item stem is combined with suitable response options, two types of instructions can be applied to SJA responses:

- *"Should"* instructions require the participant to declare their understanding of the correct or most appropriate responses to a situation. Examples of "should" instructions include "Choose the most effective/appropriate response"; "Choose the most and the least effective/appropriate responses"; "Select the three most effective/appropriate responses"; "Rank/rate the responses in terms of how effective/appropriate they are."
- *"Would"* instructions ask the participant to indicate how they would typically behave in a situation. Examples of "would" instructions include "Which of the following responses best typifies how you would respond to this situation"; "Choose the response that best reflects how you would respond to this situation and the one that least reflects how you would respond"; "Rank/rate the responses in terms of how you would tend to respond to a situation like this."

Both offer value but serve different purposes. In high-stake applications, "should" instructions serve to test the participant's understanding of how best to resolve a situation and are generally used when there are more effective or appropriate responses that demonstrates competent work

behavior. "Would" instructions serve to provide insight into a participant's behavioral style and preferences and are more appropriate in lower-stakes applications such as RJPs and surveys or when used as learner diagnostics or as training materials.

- **Response Option Formats**

  Let's explore five widely used response option formats for SJA items using our example item. These formats can be used with either "should" or "would" instructions — though we will continue with our example, set mainly in the format of "should" instructions for the sake of consistency.

  These different formats offer different levels of granularity in SJA item scoring. There is a tradeoff between greater granularity and the added complexity of the response formats that more granular scoring options require. We present five scoring options to allow the reader the opportunity to consider which formats are more appropriate for their testing and assessment programs, and which are more appropriate for the populations to be assessed using SJAs.

- **Select Most.** This straightforward response format requires the participant to select which option is the most effective or appropriate (or that would typify how they would behave for "would" instructions). Here is our example item using this "multiple choice" format for which the correct answer is D:

You are looking after Mr. Kucera who has previously been treated for prostate carcinoma. Preliminary investigations are strongly suggestive of a recurrence. As you finish taking blood from a neighboring patient, Mr. Kucera leans across and says, "Tell me honestly, is my cancer back?"

Select the most appropriate response to this situation:

a. Explain to Mr. Kucera that it is likely that his cancer has come back
b. Reassure Mr. Kucera that he will be fine
c. Explain to Mr. Kucera that you do not have all the test results, but you will speak to him as soon as you do
d. Inform Mr. Kucera that you will chase up the results of his tests and ask one of your senior colleagues to discuss them with him
e. Invite Mr. Kucera to join you and a senior nurse in a quiet room, and get a colleague to hold your 'bleep' and then explore his fears

The usual way to score this format would be to give an item score of 1 for selecting the correct answer (D) and 0 for selecting any of the alternative response options.

- **Select Most and Least.** This format offers finer granularity in assessing a participant's understanding of the demands of a situation and effective or appropriate responses by requiring the participant to select the most and the least effective/appropriate of the response options offered; the participant is required to make two responses under this format. Here is our example, in which the correct responses are to select D as Most and B as Least:

You are looking after Mr. Kucera who has previously been treated for prostate carcinoma. Preliminary investigations are strongly suggestive of a recurrence. As you finish taking blood from a neighboring patient, Mr. Kucera leans across and says, "Tell me honestly, is my cancer back?"

Select the most and the least appropriate response to this situation:

a. Explain to Mr. Kucera that it is likely that his cancer has come back
b. Reassure Mr. Kucera that he will be fine
c. Explain to Mr. Kucera that you do not have all the test results, but you will speak to him as soon as you do
d. Inform Mr. Kucera that you will chase up the results of his tests and ask one of your senior colleagues to discuss them with him
e. Invite Mr. Kucera to join you and a senior nurse in a quiet room, and get a colleague to hold your 'bleep' and then explore his fears

The usual way to score this format would be to give one point for getting Most as D and one point for getting Least as B, giving a score range between 0 and 2 for this format. An alternative approach would be to avoid partial credit and give a score only if both answers are correct.

- **Multiple Effective/Appropriate Selections.** Whereas the formats described to this point can be applied to items with at least three response options, items with more than three response options offer the opportunity to apply a "multiple response" format requiring the participant to select a number of response options that they consider the most effective or appropriate (or the most typical under "would" instructions). Here is our example item using this format. The participant is asked to select the two most appropriate responses[22].

---

[22] In the medical program from which the example item is taken, items using this format offer six response options and the candidate is asked to select the three most appropriate responses.

The maximum score for this item would be achieved by selecting C and D:

You are looking after Mr. Kucera who has previously been treated for prostate carcinoma. Preliminary investigations are strongly suggestive of a recurrence. As you finish taking blood from a neighboring patient, Mr. Kucera leans across and says, "Tell me honestly, is my cancer back?"

Select the two most appropriate responses to this situation:

a. Explain to Mr. Kucera that it is likely that his cancer has come back
b. Reassure Mr. Kucera that he will be fine
c. Explain to Mr. Kucera that you do not have all the test results, but you will speak to him as soon as you do
d. Inform Mr. Kucera that you will chase up the results of his tests and ask one of your senior colleagues to discuss them with him
e. Invite Mr. Kucera to join you and a senior nurse in a quiet room, and get a colleague to hold your 'bleep' and then explore his fears

The usual way to score this question would be to assign a score of 1 to correct responses and 0 to the other response options. For example, a participant selecting C and D as the two most appropriate responses would obtain a score of 2 for this item. Selecting C and B would obtain a score of 1. Selecting D and A would also achieve a score of 1 for this item, while selecting A and B would yield a score of 0. This format offers a score range per item of 0 (minimum possible) to 2 (maximum possible). An alternative approach would be to avoid giving partial credit and give a score only if both answers are correct.

- **Rank the Response Options**. This format requires the participant to rank the response options from the most to the least effective or appropriate response. This format requires evidence that there is a clear and unambiguous order to the response options. If such an unambiguous order can be obtained, the advantage of this item is that all response options can be incorporated in the scoring of an SJA item.

Here is our example using the format:

> You are looking after Mr. Kucera who has previously been treated for prostate carcinoma. Preliminary investigations are strongly suggestive of a recurrence. As you finish taking blood from a neighboring patient, Mr. Kucera leans across and says, "Tell me honestly, is my cancer back?"
>
> Rank the appropriateness of the following actions in response to this situation:
>
> a. Explain to Mr. Kucera that it is likely that his cancer has come back
> b. Reassure Mr. Kucera that he will be fine
> c. Explain to Mr. Kucera that you do not have all the test results, but you will speak to him as soon as you do
> d. Inform Mr. Kucera that you will chase up the results of his tests and ask one of your senior colleagues to discuss them with him
> e. Invite Mr. Kucera to join you and a senior nurse in a quiet room, and get a colleague to hold your 'bleep' and then explore his fears

The ideal ranking is DCEAB. One way to score this question would be to give a score of 1 if the ranking is correct, or 0 otherwise. Or it's possible to use more complex scoring, giving credit for where each option is placed.

- **Rate Each Response Option**. A variant of the ranking approach is to require the participant to provide a rating of each response option in terms of effectiveness or appropriateness (or typicality for "would" instructions). Here is our example using that format:

You are looking after Mr. Kucera who has previously been treated for prostate carcinoma. Preliminary investigations are strongly suggestive of a recurrence. As you finish taking blood from a neighboring patient, Mr. Kucera leans across and says, "Tell me honestly, is my cancer back?"

Rate each option in terms of how appropriate it is in responding to the situation.

Enter: 1 for an option you think is **Highly Inappropriate**

2 for an option you think is **Inappropriate**

3 for an option you think is **Neither Appropriate nor Inappropriate**

4 for an option you think is **Appropriate**

5 for an option you think is **Highly Appropriate**

a. Explain to Mr. Kucera that it is likely that his cancer has come back
b. Reassure Mr. Kucera that he will be fine
c. Explain to Mr. Kucera that you do not have all the test results, but you will speak to him as soon as you do
d. Inform Mr. Kucera that you will chase up the results of his tests and ask one of your senior colleagues to discuss them with him
e. Invite Mr. Kucera to join you and a senior nurse in a quiet room, and get a colleague to hold your 'bleep' and then explore his fears

A correct response string using this format might be to rate A as 1, B as 2, C as 4, D as 5 and E as 3. The question could be scored with one point per correct answer, or with weightings also given for close answers.

The five response formats described above are among the most popular used with SJAs. Each comes with a benefit such as the simplicity of the Select Most format and the granularity of scoring offered by other formats. Each also has potential costs in terms of the comprehension required of participants (and therefore the participant experience) and the complexity of scoring algorithms. Response formats are key choices in constructing SJAs that require consideration of the target populations the SJA will be administered to. If you are new to SJAs, we recommend the Select Most and Least option for your initial consideration.

If questions with different response formats are mixed together, it may be necessary to weight some items to ensure all are equal. For example, if mixing Select Most and Select Most and Least questions in the same assessment, you may want to allocate two points to the correct answer in Select Most questions to weight the questions equally. Please note that there are many other approaches to response format and scoring and we have concentrated on a few for brevity. The publications referenced in this paper provide more sophisticated options.

Generally, overall scores are obtained from SJAs by simply summing item scores. How many items are required to construct a reliable and valid SJA score will depend on the quality of the items, but this leads to a frequently asked question: how long should an SJA be?

Length and timings for SJA assessments

The question of how long an SJA assessment should be arises fairly frequently with those new to SJA content. The answer may, however, be slightly unsatisfactory. The honest answer is that it depends.

In pre-hire settings using RJPs, five or six items may be sufficient for an SJA assessment. In the higher-stakes scenarios when SJA questionnaires are used to inform hiring (employee selection) decisions, they may be as short as 14 to 20 items when used alongside other assessments[23]. For the UK Foundation Programme, which uses two different response formats for its SJA items, SJA exams contain 70 SJA items. These exams contain 60 scored SJA items and 10 pilot items to allow the program to continuously refresh SJA item pools. The length of this exam reflects the high-stakes nature of the SJA application as well as the weight it carries in overall decisions to admit or reject participants applying for the medical program.

---

[23] An example is the use of SJA questionnaire by the European Personnel Selection Office (EPSO). Examples of their use of SJA questionnaires can be found via https://epso.europa.eu/how-to-apply/sample-tests_en.

While the length of SJA assessments vary, there is a broadly consistent benchmark in terms of timings for them. Typical average response times for SJA items vary from one to two minutes depending on the complexity of those items. These average timings provide a rule of thumb that can be used for planning the administration times for SJA assessments.

For a 20-item SJA questionnaire, that rule of thumb would suggest a total administration time of between 20 and 40 minutes. These are indicative timings that should be validated through pilot trials and analysis of item response times and overall completion times of the full SJA assessment.

Defending the security of SJAs

As with any other higher-stakes assessment, security of item content will be a concern. In addition to protocols for permissions to access SJA items and systems for managing access and the usual security procedures like proctoring or using a secure browser such as Questionmark Secure, there are some simple design steps that can strengthen the security of SJA content:

- **Use "should" instructions.** Research shows that "should" instructions are less susceptible to coaching and faking primarily due to the fact that they require identification of correct, better or ideal responses. This framing of an SJA item requires the participant to apply reasoning to solve the item and renders items less susceptible to the influence of tendencies toward self-deception or impression management (such as faking good responses)[24].

24 McDaniel, M. A., Hartman, N. S., Whetzel, D. L., and Grubb, W. L. (2007) Situational judgement tests, response instructions and validity: A meta-analysis. Personnel Psychology, 60, 63-91. 25 When the candidate is required to select most and least options, the correct response depends on identifying two options in order. With four response options, this gives twelve permutations of possible answers involving pairs of options and therefore a probability of 1 in 12.

- **Use formats involving the selection of more than one response.** For an item with four response options, the probability of identifying the correct answer by chance alone is 1 in 4 or 0.25. By requiring the participant to identify most and least options, that probability reduces by a third to 1 in 12 or 0.08[25]. While items allowing for multiple selections among response options do place higher comprehension demands on participants, this can be managed by ensuring simplicity in SJA item instructions and providing example items explaining what is required of participants in responding to items.
- **Use randomisation or shuffling of item response options.** If your technology platform permits, it's useful to randomize the ordering of response options so that different participants are presented with response options in different orders.
- **Use alternate forms or selection of items from SJA item pools.** Alternate forms (instances of an assessment) can be deployed to minimise the exposure of SJA content or to preserve scoring keys where the order of answers is varied between different forms of the SJA. An alternative is to create an SJA item pool and draw on that pool by randomly selecting items to construct an assessment that is different each time it is administered to a participant.

Pitfalls to avoid

To conclude this section, we offer a brief list of pitfalls to avoid and some to-dos in developing and deploying SJAs.

- **Avoid trying to cover everything in your SJA**. SJAs function best when there is clarity about what is being measured and when that measurement is focused on a limited set of dimensions (skills, competencies or compliances). Four to five dimensions covered by at least four items per dimension is a good starting point to ensure that your SJA assessments have focus and will yield stronger psychometric evidence supporting SJA scores.

- **Avoid simply describing situations in the item stem**. Measuring judgment requires a choice to be made. SJA items work best when there is a dilemma that is clearly linked to the need for participants to demonstrate a clearly defined domain of knowledge, skill, competency or compliance in resolving the situation presented in the item stem.

- **Avoid response options that are disconnected to the situation in the item stem.** Item stems and response options should function as a coherent whole in tapping into judgments linked to a clearly defined knowledge, skill, competency or compliance required to resolve the situation.

- **Avoid response options that vary significantly in length. Participants** look for clues to the correct or desired response to an SJA item and may interpret shorter or longer response options as correct or preferred answers when those options may be neither correct nor preferred.

- **Do check scoring keys with SMEs.** As well as avoiding wasted effort and cost in piloting and tryouts of SJA items and assessments, identifying where there is a lack of consensus on most and least effective/appropriate is an essential step in ensuring that SJA content is functioning as expected. This also provides insight on how to improve SJA content design.

- **Do check that your scoring of SJA items and questionnaires is defensible.** SME checks also provide valuable evidence that defends the rationale behind SJA items and assessments. In highstakes applications, evidence supporting defensibility is further strengthened by data showing that the SJAs do identify those achieving desired outcomes (e.g., competent or higher performance).

- **Do check SJA content for cultural relevance and sensitivity.** We all carry our own cultural lenses, so it is sensible to check that SJAs reflect the diversity of the populations they will be administered to, particularly in multinational settings. Sensitivity checks are especially important when SJAs are to be deployed in different languages. Conducted early, these checks flag items that may need to be adapted in the originating language for them to function in another language.

- **Do check the ongoing relevance of your SJA content as circumstances change.** The needs of organizations and what defines competent performance are continuously changing. It is sensible to carry out periodic and regular reviews to check that SJA content retains relevance and credibility.

Having described some of the key steps for constructing SJA items and assessments, let's explore what to look for in technologies to support SJA content development and deployment, as well as content maintenance and security.

5. Leveraging technology to maximize the value of SJAs

Previous sections of this white paper have explained where SJAs can be useful and how to identify suitable dilemmas and questions around them. Most users of SJAs will seek to deploy them with technology. This chapter gives some guidelines on where assessment management technology is useful to deploy SJAs and what capabilities to look for in selecting appropriate tools.

It's helpful to think about six key stages in assessment development as follows:

1. **Planning assessment** – the steps to blueprint and organize your assessment before you develop the items
2. **Authoring items** – creating the individual SJA items
3. **Assembling assessment** – putting the SJA items (alongside any other items) into the test, quiz, survey or exam
4. **Pilot and review** – checking the questions and assessment prior to production delivery
5. **Delivery** – presenting the SJAs to participants, gathering the answers and scoring them
6. **Analyze results** – obtaining reports to help feedback to improve the process and to share with stakeholders.

These steps mirror the considerations shared earlier in building effective SJAs. Using these steps — from Planning assessment to Analyze results — we will now share technology suggestions (useful to consider with any assessment management and deployment system) and specific ways that Questionmark technology can help in designing and administering SJA assessments as well as managing their content. These capabilities are available in Questionmark's SaaS solution Questionmark OnDemand, in our behind-the-firewall solution

Questionmark OnPremise and in SAP Assessment Management by Questionmark.

Planning Assessment

Key capabilities in defining SJA requirements	How Questionmark technology can help
1. **Consider a job task analysis (JTA) survey to identify the facets of workplace judgment to assess.** A JTA survey can ask job experts what tasks are performed in a job, how often they are performed and how important they are. You can also use such a survey to ask what "softer skill" attributes are important for the job and need testing for, as well as to identify possible compliance challenges. A JTA will allow you to build a table of attributes such as the example for medical doctors we shared in Section 2.	Questionmark has comprehensive features for creating, delivering and reporting on JTA surveys, including a special JTA question type and JTA-specific reports like the one shown below for presenting information from job task analysis surveys.
2. **Survey experts to obtain examples of real-life dilemmas and/or to rate possible dilemmas for utility in SJAs.** It is essential to be able to capture, manage and evaluate examples of situations as well as effective versus less effective response options.	Questionmark has comprehensive capabilities for authoring, delivering and reporting on surveys. Survey reporting allows you to filter results by demographics and dates and includes frequency analysis graphs. You can also download data in Excel format to do custom analysis.
3. **Identify topics as measurement attributes to build your assessment on.**	Questionmark provides comprehensive item bank facilities, allowing use of topics and metatags to organize items. It also permits selection of items by topic and

For most applications of SJAs, it is useful to consider a (usually hierarchical) topic structure. Topics are used to organize items, ensuring appropriate coverage of topic areas within an assessment, and to present sub-scores by topic to aid coaching and identify areas that need improvement or present compliance challenges.

the presentation of topic scores to participants and within reports. Another advantage of using an item bank is that it makes questions easier to re-use in other assessments.

Authoring Items

Key capabilities in designing SJA content	How Questionmark technology can help
1. **Most importantly, you need a system that can deliver the presentation of the SJA items you have selected.** We regard these five question types as the most useful: • Select Most • Select Most and Least • Multiple selections • Rank response options • Rate response options	Questionmark allows you to create questions for all five of these types: • "Select Most" questions are multiple choice questions within Questionmark software. • "Select Most and Least" questions can be delivered as Matching questions. See the screenshot below. • Questionmark's ranking question type allows "Rank response options." • The "Rate response options" can be delivered in Questionmark using the Survey Matrix question type (for surveys) or Pull-down list or Knowledge Matrix for tests.
2. **You also need technology that can score the items in the ways you need.**	Questionmark has a wizard approach to creating questions. It allows simple right/wrong or score per choice scoring and also has an Advanced Editor that

allows more complex scoring. With the Advanced Editor, you can assign a score to any combination of response selection and can make complex SJA partial scoring if desired.

3. SJAs can benefit from graphics, video or other media to provide appropriate context for a dilemma.	Questionmark makes it easy to add media into question wording, so you can provide resources that add context to a dilemma.
4. You need a collaborative environment where authors and reviewers can cooperate to build SJAs. Developing SJAs is usually a team effort. You usually need to iterate on a question to get it to a suitable measure of judgment. Recording item development history is very useful for showing legal defensibility. You will want to record item changes, be able to roll back to previous versions, control access to content and see item history.	Questionmark gives each authoring user a separate login, and users can have different roles and permissions, for example permitting reviewers access to certain topics or read-only rights. All items are versioned, and authors enter comments when an item changes. You can then roll back to previous versions and see a "track changes" style view of item history as in the example below.
5. Tryouts of the look and feel of SJA items are valuable to guide the iterations toward the final format and participant experience of SJA items. This allows you to put yourself in the state of mind of a participant using judgment.	Questionmark lets you try out a question or assessment at any time while developing it. ▶ Try out
6. The capability to shuffle the order of response options each time the item is delivered is a powerful method for	Questionmark permits you to shuffle options in several question types.
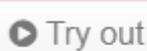
defending the security of SJA content and answer keys.

This capability can also enhance assessment by minimizing order effects as well as biases that may arise from the authoring of content.

The example shown below illustrates authoring of an SJA item using the Select Most and Least response format in Questionmark OnDemand. Further examples of how content can be authored within Questionmark are provided at **www.questionmark.com/go/example-sja**.

Assembling assessment

| Key capabilities in constructing SJA assessments	How Questionmark technology can help
1. **Needs may be for a standalone SJA in the form of several SJA items or for a blend of SJA and knowledge-based items (or other content).**	Questionmark treats SJA questions in the same way as other questions, and it's easy to construct • SJA only assessments • Assessments with a block of knowledge questions followed by a block of SJA questions • Assessments that blend question types where SJA questions and other questions are interspersed.

2. **Random selecting and ordering of questions can be helpful.**
It can be useful to
• Choose question order randomly to avoid ordering effects and as an anti-cheating measure
• Select questions randomly from an item bank to reduce question exposure and as an anti-cheating measure
• Select different questions on subsequent attempts from those a previous participant saw previously
	As shown in the screenshots below, Questionmark allows you to select questions within an assessment
	• Individually
	• By topic and/or sub-topic
	• At random from a topic
	• At random (but excluding those delivered previously to the same participant)
	
	Select questions to include
	○ All questions from topic
	  ☐ As single questions
	◉ Randomly select [2] from topic
	  ☐ Avoid previously delivered
	○ Single question
	☐ Include sub-topics
	☐ Filter questions by Metatags ✚ Add
	
	• By metatag: Filtering items by metatag enables authors to

further refine item selection from specified topics and subtopics.
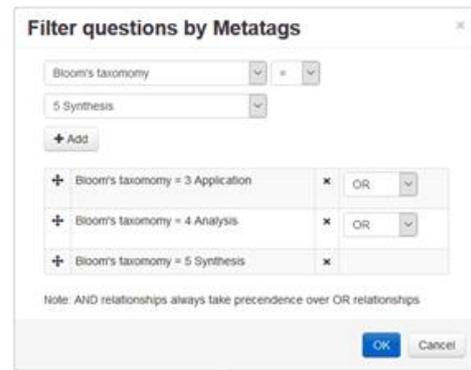
3. Allow a survey to be presented anonymously.

When presenting SJAs in a survey, for example in regulatory compliance where you are seeking to assess opinion or likely behavior, it can be useful to make results anonymous. This encourages participants to answer more freely, particularly on sensitive matters.

Questionmark permits surveys with anonymous results. A participant is authenticated to log in to an SJA assessment, but their name and personal details do not get saved.
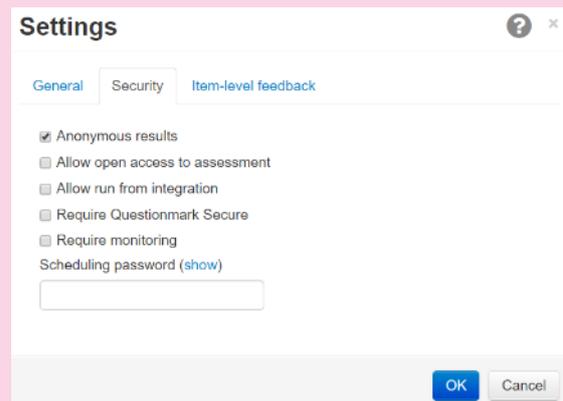
4. Make it easy to manage the translation of SJAs.

Because SJAs involve complex thinking, it is important to present them in a language the participant is very familiar with — so translation is often required. Ensure that you review and pilot all translations to avoid loss of fidelity due to cultural issues or translation errors.

Questionmark has a capable Translation Management System which allows you to export question text for translation, then reimport it for review and piloting.
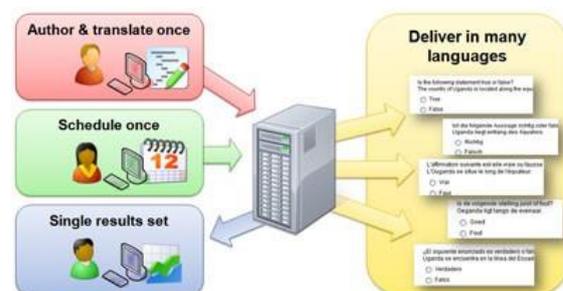
5. Review the item bank and update over time. Work conditions change, so it's important to manage your item bank as the job requirements change. You need to regularly review the item bank and change or remove items that are no longer appropriate.	With Questionmark, you can search an item bank (e.g., to look for all items that mention a key phrase) and mark items as retired — so they stay in the item bank for historical and statistical purposes but are not selected for assessment delivery.

Pilot and review

| Key capabilities Pilot and Review	How Questionmark technology can help
1. Pilot the questions with the same technology and user interface you will deliver them with. This makes it more likely that the data gathered in the pilot will validly predict production performance.	With Questionmark technology, you can easily publish an assessment for pilot delivery. You can use a different authentication method for a pilot to that used in production delivery. For example, you could construct individual usernames/passwords to use for a pilot, even if you are delivering the production assessment via SAML or an LMS.
2. When your program is ongoing, it can be useful to include pilot or beta items within production assessments on an experimental basis. This lets you gather item statistics with real participants but avoid biasing scores with unproven questions. Once the items have been validated, you can then remove the experimental status and use the items "for real."	In Questionmark, you can mark items as experimental. This means they gather data used for item statistics but not for scoring the assessment. Question delivery status Normal Normal Retired Incomplete Experimental Beta
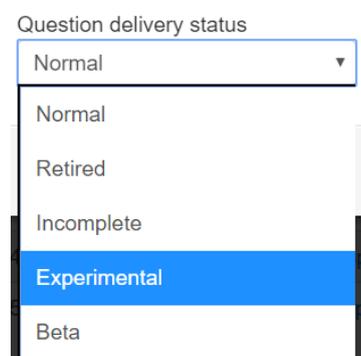
3. **Gather qualitative comments from your pilot participants to help improve questions.**

SJAs rely on communicating the dilemma and the context around it effectively so that the participant can exercise judgment. Encouraging comments during the pilot allows participants to advise on ambiguity or missing assumptions.

Questionmark allows you to add a comments box to questions to allow participants to offer feedback. You can then review their comments when you evaluate the pilot and remove the boxes from question presentation once the items are validated.

Add comment box

4. **Obtain quantitative item statistics from the pilot to validate and improve questions.**

Item statistics give crucial quantitative data to determine if your SJAs are appropriate for production use. You will significantly improve the validity and reliability of your SJAs if you conduct item analysis on the pilot and adjust or remove questions based on the results.[26]

Questionmark has a comprehensive item statistics report that can easily be configured to produce a report on a set of pilot results.
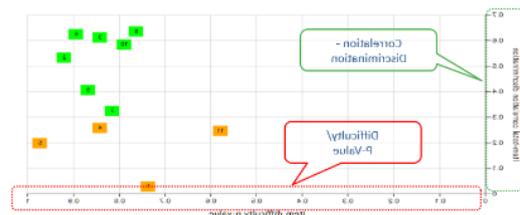
For example, in the example item analysis above, you would want to look at the questions in amber to investigate why there is low item correlation discrimination.

---

26 For more information on item statistics, see Questionmark's eBook on Item Analysis Analytics — available at **www.questionmark.com/resources/whitepapers**.

Delivery (deployment of SJAs)

Key capabilities	How Questionmark technology can help
1. **You need to be able to integrate SJA assessments into your wider communication mechanisms — whether LMSs, applicant tracking or other systems.**	Questionmark has interfaces to many systems out of the box and simple-to-use APIs that allow it to connect to almost any IT system. These include support for SAML, support of calls from LMSs via AICC and SCORM, interfaces with SAP software and Office 365, and SOAP/REST APIs.
2. **You need a compelling, robust and easy-to-use participant UI that works on all devices including mobile devices.**	Questionmark uses responsive design to ensure wide browser support and permit assessment delivery on mobile devices. A wide range of delivery templates and styles are readily available.
3. **You need a system to instantly score the SJA results, and it's often useful to present feedback to the participant.**	Questionmark stores SJA answers in a secure database and scores them according to the rules set up in authoring. You can optionally deliver feedback at three levels:

	• Assessment feedback (e.g., pass/fail) • Topic feedback (scores and feedback on a topic basis indicating areas of improvement) • Question feedback
4. You need a scalable and robust system.	Questionmark OnDemand delivers over 10 million assessments per year and can easily deliver large numbers of SJA assessments robustly. Questionmark's Save As You Go (SAYG) capability ensures that if an assessment is interrupted by a device failure or connectivity issue, the assessment can easily be resumed on another device from the last question taken.
5. You need security against cheating, hacking and other breaches of confidentiality or integrity.	Questionmark software has many "anti-cheating" capabilities including the ability to monitor or proctor participants and a secure browser, Questionmark Secure, which helps exam integrity. Questionmark itself is ISO 27001 certified; see **www.questionmark.com/trust** for more about Questionmark's security.
6. Provide practice tests. To familiarize participants with SJA formats, it's sensible to provide some practice items they can take prior to taking the real assessment.	It's easy to construct and deploy practice tests in Questionmark. One option is to use the open access entry point that allows them to be hosted without authentication.
7. You need a system that will meet your data protection concerns (including GDPR if applicable).	Questionmark meets most data protection requirements including GDPR and FERPA. See also **www.questionmark.com/trust**.

Analyze results

Key capabilities in analyzing data from SJAs	How Questionmark technology can help
1. You need comprehensive reports to allow professional decisions to be made as a result of SJAs. Useful reports include reports on individuals and groups, filtered by organizational demographics. You will want to know pass/fail, topic scores and question results. When SJAs are used as surveys, you will want to be able to get meaningful analysis of the data. Once you have established a baseline, reports showing trends (i.e., change over time) are hugely valuable.	Questionmark has very strong reporting capabilities that let you report on assessments in a very powerful way. There are also ways to connect to results via APIs for more advanced analytics. See **www.questionmark.com/content/reporting-analytics** for full details.
2. You need a report to share with participants. When using SJAs to measure competence, it's often helpful to create a coaching report, which contains the questions, the answers and feedback. This can either be used directly by the participant for self-improvement or by the participant's manager or another coach.	Questionmark makes available a configurable coaching report suitable for handing back to the participant to help them understand how to improve, or for their manager to run through with them. Coaching reports can give topic feedback to help identify weaknesses in specific areas and/or give answers and feedback on specific questions.
3. You need to run item analysis regularly. As you build up more data from delivery, it's important to run item analysis reports	As described above, Questionmark has a very effective item analysis report.

regularly. Weaknesses in questions may become more apparent when there is a larger volume of results. Question effectiveness can also change if a question is exposed — for example if an SJA dilemma is widely discussed within your organization, it may no longer be useful to include in an assessment.

4. Run a report to calculate reliability.

The reliability of an assessment is usually measured by a statistic called the Cronbach's Alpha, and it's sensible to review this to check the consistency and reliability of your assessment.

Questionmark's test analysis report calculates Cronbach's Alpha and other useful test statistics.

Standard error of mean	0.09/11 (0.82%)
Standard error of measurement	1.16/11 (10.55%)
Skew	-1.322
Kurtosis	1.025
Test reliability (Cronbach's Alpha)	0.835

5. Ensure your system can track time taken.

To confirm you have allowed participants enough time, you need to be able to track the time taken on an assessment and the time taken by participants on each item.

Questionmark tracks time taken on each assessment and (providing you use the question-by-question presentation format) for each question.

.

6. Conclusion

The impetus for this paper was a simple question: with judgment in the workplace under ever greater scrutiny and with SJAs gaining ground to meet many employment needs, is there wider value in the application of SJAs to meet test, exam and assessment needs? We thought there was and offer this paper to help you evaluate how SJAs can enhance your programs. In summary, we have identified these key arguments in favor of considering SJAs:

1. Judgment is a key competence in many job roles. It is useful to narrow down specific dimensions of judgment important in job roles via job task analysis or similar techniques. If you are assessing pre-hire, post-hire or for certification or compliance purposes, then measuring judgment will usually improve the accuracy and utility of your assessments.

2. SJAs are a way to measure judgment by presenting a dilemma to participants and having them rate options to resolve it. In comparison to simulations and role-playing exercises, SJAs are often more objective and more cost-effective, so they can provide valid and reliable — and thus trustworthy — ways of measuring judgment in the workplace. SJAs also allow assessments to operate at higher levels of the Blooms Taxonomy.

3. The key to authoring an effective SJA is to identify a dilemma that requires a choice to be made and where using judgment to resolve that choice is clearly linked to a needed domain of knowledge, skill or competency in the job role. Because an SJA is only useful if it requires professional judgment to answer well, you must involve Subject Matter Experts in authoring and reviewing SJAs.

4. There are many options and formats for SJAs. One of the most effective is the Select Most and Least approach, where 4 or more options are presented and the participant must identify the most and least effective or appropriate. This has the advantage of being relatively simple for the participant to consider, while making it but much harder to guess the right answer than it would be to select a single most appropriate response.

5. Questionmark technology can significantly aid authoring, delivering and reporting on SJAs. Questionmark supports the question types and scoring

options needed for SJAs and provides an environment in which it's easy to manage the engagement with SMEs that is needed to author them successfully. It also provides effective ways to pilot SJAs and gather qualitative and quantitative feedback to improve the items. Questionmark software provides a reliable, robust, scalable and effective means to deliver SJAs and produce effective reports on them.

6. We encourage you to consider how you could deploy SJAs within your assessments:

    - Within regulatory compliance, SJAs can be used to assess competence and survey how participants think personnel will respond.

    - For certification programs, SJAs provide a relatively inexpensive way of extending coverage to higher-order and softer skills required by participants.

    - For pre-hire assessments, SJAs provide a way for candidates and organizations to get insight on a candidate's capability and fit for the role.

    - And for post-hire training and workforce development, SJAs give the ability to measure important workforce competencies not easy to assess in other ways.

About Questionmark

Questionmark assessment and portal solutions enable organizations to measure knowledge, skills and attitudes for certification, channel expertise, workforce learning and regulatory compliance. Questionmark's assessment management system, available as a cloud-based solution or for on-premise deployment, enables collaborative, multilingual authoring; multiple delivery options including mobile devices; trustable results and comprehensive analytics.

Complete details are available at https://www.questionmark.com

Questionmark
35 Nutmeg Drive, Suite 330
Trumbull, CT 06611
United of States of America
Tel: (800) 863-3950
Fax: (800) 339-3944
info@questionmark.com

Questionmark
Moor Place, 1 Fore Street
London EC2Y 9DT
United Kingdom
Tel: +44 (0)20 7263 7575
Fax: +44 (0)20 7263 7555
info@questionmark.com