

IRIS

Situational Judgement Tests

User Guide



Copyright © 2012 Pearson Education Ltd or its affiliate(s).
All right reserved.

Pearson, the **Pearson** logo and **TalentLens** are trademarks in the U.S. and/or other countries of Pearson Education, Inc, or its affiliate(s).

Printed in the United Kingdom

www.talentlens.co.uk

Acknowledgements

The IRIS measurement context and the tests themselves were developed by the following people:

Helen Baron

Chris Lewis

Alice Revell

Abi Scott

Neil Scott

Paul Smith

Paul Yarker

Many other people contributed to the development of the IRIS tests and we would like to thank all those who contributed ideas or their time to write or review material, participate in focus groups, complete a test, or otherwise help in their development. In particular we would like to thank: Sue Overy, Wyn Davies, Matt Stewart and Alie Carvey.

We are very grateful to the representatives of many organisations who made up the teams of subject matter experts to ensure the tests reflected the situations and scenarios typically encountered within the roles.

Foreword

This user guide describes the IRIS tests and provides guidance on their use as well as a summary of technical information pertaining to the tests. Users who are interested in more detailed technical information or details of the test development process for IRIS should contact the TalentLens team on 020 7010 2866.

Contents

Acknowledgements	i
Foreword	ii
Chapter 1: About IRIS	
Overview of the tests, what they measure and the range of IRIS test available	
Situational Judgement Approach	1
What is being assessed?	4
Use of IRIS tests	7
Chapter 2: Using IRIS	
Administration of the tests	
Scoring and norms	9
Preparing candidates for testing	9
Chapter 3: Results and Interpretation	
IRIS scores, their interpretation and the different outputs available	
Using scores	10
Interpretation	12
Customer Service IRIS Feedback report - 2 page sample	14
Graduate IRIS Feedback Report - 2 page sample	15
Sales IRIS Feedback Report – 2 page sample	16
Chapter 4: Technical Information	
Summary of the technical information for the IRIS tests	
Development	17
Scoring	18
Reliability	18
Validity	18
Content Validity	18
Face Validity	19
Criterion Validity	19
Group Comparisons	19
Figures	
Figure 1: Example Question {Customer Service test}	3
Tables	
Table 1: Competencies and behavioural areas measured by each test	4
Table 2: Task categories for the Customer Service IRIS test	5
Table 3: Task categories for the Graduate IRIS test	5
Table 4: Task categories for the Sales IRIS test	6
References	20

Chapter 1: About IRIS

The IRIS tests are online measures of work suitability designed to be relevant to many different roles and industry sectors. The tests were carefully developed and offer a ready-made solution to measuring job suitability for short listing and other assessment needs. IRIS tests use a situational judgement approach and respondents are asked to determine how appropriate different responses are to a given scenario. The test content consists of relevant situations and scenarios and can be tailored to different organisational and role requirements. Users can choose tests suitable for different job levels, content and contexts. There are currently three IRIS tests available relating to the following roles:

Applicable Roles	
Customer Service	Customer service Entry-level roles Administrative
Graduate / Junior Manager	Graduate level Junior management
Sales	Sales roles focussing primarily on new business (e.g. telesales)

Recent research (McDaniel et al. 2001) has shown that the situational judgement tests are very effective predictors of job performance in a variety of different situations. Typically situational judgement tests are designed for one particular job and are specific to that role and organisation. The IRIS tests are different as they have been designed to a generic competency framework and assess those aspects of performance which are common within job families across organisations. This means that test users can gain the benefit of this approach without the expense of developing their own instrument.

Situational Judgement Approach

While measures of ability have long been proven as successful predictors of job performance (e.g. Robertson and Smith, 2001) they only relate to one, albeit important, aspect of performance. Research into job performance has shown that there are many other aspects beyond technical ability (e.g. Motowidlo and Van Scotter, 1994). The personal style and the approach of the individual can be just as, or sometimes even more, important in determining the quality of overall performance. Competency models typically include many dimensions which are only weakly related to thinking ability. It is addressing these aspects of performance where the situational judgement methodology

can come into its own. While interviews, role play exercises and suchlike can be effective measures of such aspects of performance they are expensive to use, requiring time of candidates and highly trained assessors. The situational judgement approach provides an assessment for which both administration and scoring can be automated. It is therefore highly suitable for use as a short-listing tool in selection where large numbers of candidates need to be assessed. It can also be useful where time constraints or the absence of appropriately trained staff mean that other types of exercise cannot be used.

The situational judgement test assesses the candidates' approach to work, the way they deal with different circumstances, people and dilemmas. In the IRIS tests candidates are presented with a context description reflecting real life job circumstances followed by a number of realistic scenarios that might occur in the given context. They are asked to rate a number of potential actions in response to the scenario according to how effective they would be in dealing with the situation. An example of this is provided in figure 1 overleaf.

The tests typically take between 20 and 45 minutes for the candidate to complete, but they are untimed. The tests are administered online and administration is controlled as candidates can only access the tests if they have a login supplied by an employer registered to use the tests. Supervised administration is also possible anywhere where a computer with a browser linked to the internet is available. However, most candidates will complete the tests unsupervised on-line at their own convenience. The tests are available 24/7 within the period of testing specified by the user. They are supported by a dedicated helpdesk in case of queries from candidates about the tests. This allows users to manage the assessment of large numbers of applicants very easily.

Research suggests that candidates often have more positive perceptions of situational judgement tests than traditional recruitment methods as they immediately recognise the job-relevance of SJTs and why they are being used in the assessment process. As a result of this, their use can enhance the employer brand. The realistic nature of the scenarios means that candidates also receive a realistic preview of what a job might entail while completing the test. SJTs allow candidates to decide whether the job is in line with their expectations and if they wish to continue with the application process, meaning they can promote effective self-selection among candidates.

Figure 1: Example Question (Customer Service test)

Background

You work for the Senior Concessions Agency (SCA).

The SCA is responsible for the administration and operation of non-means tested and tax-free concessions for the over 60s. At present this covers winter heating allowance, travel and parking permits and, recently introduced, essential food vouchers'.

The SCA operates on a regional basis. You work in the team, for the East Anglian Region, that deals with travel and parking permits. Unlike the other concessions, people must apply for their travel and parking permits if they require them when they reach the age of 60.

The SCA writes to people shortly before their 60th birthday informing them of the concession, and enclosing the necessary application forms. Your team handle queries and complaints from the public, send out information and application forms, and send out permits, for which a photograph is required. Permits are valid for four years.

Work is allocated by the Team Leader on a daily basis as necessary.

Scenario

Information letters, with enclosed application forms for travel and parking permits, are despatched by a junior member of the team every Friday afternoon. Ruth, the newest member of the team who you do not know very well, performed this duty on Friday. It is Monday morning and Ruth tells you that she thinks she forgot to enclose the application forms for as many as half of the 500 letters sent out on Friday. She asks you what she should do. It has been worrying her all weekend.

How appropriate are the following responses to the situation?
Make your choice by clicking the appropriate number.

1	2	3	4
Totally unacceptable	Not helpful	OK but not ideal	Good thing to do

A.	Advise Ruth that it is not a major problem as those who did not receive application forms are likely to contact the SCA, and forms can then be sent to them. Therefore she does not have to do anything.	1	2	3	4
B.	Complain to the Team Leader that this was an easy task to perform and Ruth's mistake will cause more work for the team.	1	2	3	4
C.	Find out from Ruth what she would like you to do to help her.	1	2	3	4
D.	Go to the Team Leader with Ruth to point out that this was an easy mistake to make.	1	2	3	4

What is being assessed?

IRIS tests measure the candidate's capability or competency by assigning each action an appropriate rating and scoring candidates higher the nearer their responses are to the assigned rating (see chapter 3 for a more detailed discussion of scoring). The candidate must exercise judgement in a broad range of areas to answer the exercises correctly.

In developing the IRIS tests an array of job competencies were considered for each of the tests to reflect the different roles. In each test, the scenarios reflect the generic job competencies in a range of employment sectors across both public and private sector settings. The competency models for each test are outlined in Table 1 below.

Table 1: Competencies and behavioural areas underpinning each test.

Customer Service	Graduate	Sales
Coping with pressure	Coping with pressure	Communication
Adapting to change and drive	Adapting to change and drive	Influencing
Energy and motivation	Energy and motivation	Customer focus
Initiating and developing relationships	Initiating and developing relationships	Initiating and developing relationships
Problem solving	Problem solving and analysis	Analysis
Team working	Managing and developing staff	Working with others
Reliability	Providing direction and decisions	Planning and organising
	Planning and organising	
	Innovation and creativity	

Situational judgement questions by their nature tend to be holistic measures. Each scenario carries with it the complexity of real life. For example, in responding to the scenario in figure 1 candidates must evaluate the different actions in terms of how well they would solve the problem by meeting the customers' needs, meeting the requirements of the organisation and managing the performance of team members. This requires a number of different competencies.

Thus overall the IRIS exercises are measuring the candidate's ability to recognise appropriate behaviour in general work related contexts. They complement cognitive ability tests because they relate predominantly to the non-cognitive elements of performance.

While the IRIS tests require candidates to read several pages of text, the overall reading level is not high and reflects the minimum level for the type of post being selected for. Customer service and sales test exercises use very straightforward language with the language for the graduate exercises a little higher level. The level of language should not create a barrier for a person with the capacity to carry out a typical target role.

The themes of the tasks in each of the different tests are given in Tables 2, 3 and 4.

Table 2: Task categories for the customer service IRIS test.

Tasks for Customer Service test:

The candidate is introduced to situations such as dealing with relationships at work, prioritising workload and making decisions under different conditions e.g. in stressful or overworked environments.

The candidate is required to face situations that involve customer complaints, dealing with managers and fellow colleagues and managing their time in a busy working environment.

This exercise requires the candidate to deal with customer complaints through face-to face and phone call interaction. It assesses customer service skills, decision-making and the appropriateness of taking ownership, working autonomously and when to involve others.

Table 3: Task categories for the graduate IRIS test.

Tasks for Graduate / Junior Manager test:

The candidate (who for example takes the role of a departmental manager) is required to make decisions based on the management of staff, reducing numbers due to a recruitment freeze and developing strategies to improve services.

The candidate (managing a small team of customer support staff) is placed in a position where staff morale is low and in situations where they may need to motivate team members, deal with their levels of performance and develop improvement strategies.

The candidate assesses whether new agreements and working practices may benefit or cause harm to the company. It forces the candidate to make decisions based on change and to consider the effect that changes may have on the team.

Table 4: Task categories for the sales IRIS test.

Tasks for Sales test:

Managing new enquiries

Candidates are presented with new queries from individuals or organisations that may have not had previous exposure to the service or the host company.

Retaining customers

Candidates are presented with existing customers who, for whatever reason, are dissatisfied with the host company's service.

Solving client problems

This task covers problems or issues that are presented by existing customers.

Navigating gatekeepers

Candidates are presented with the need to interact with individuals for who decisions on purchasing or service hiring may or may not be their responsibility.

Advancing the sale process

Candidates are expected to interact with individuals or organisations in such a way that the process of reaching a sale can be brought closer.

Prospecting / Time management

Effective time management is a key requirement for individuals working in sales. Throughout this task candidates are asked to prioritise activities effectively.

Understanding the customer

This task covers issues or problems that may be faced when prospecting for new business.

Use of IRIS tests

The IRIS exercises will be appropriate for a wide range of jobs, but primarily customer service, entry level, sales, graduate and junior management roles. This might include the assessment of graduate recruits and school leavers as well as more experienced applicants for these types of jobs. Users should satisfy themselves that the content of the exercises chosen is relevant to the role vacancies being filled. The exercises can be tailored to make them more relevant to particular industries or organisations if required.

Because the exercises can be administered via the internet they are particularly useful in short-listing from large applicant pools. The tests are fully automated and a single administrator can manage the assessment of a considerable number of candidates. Reporting is also automated which reduces the onus on the trained test user.

While the IRIS exercises are most likely to be used early in a selection procedure they can also be useful at the later stage. They provide an objective indicator of suitability which can be used to supplement the subjective view derived from an interview process. They provide a global score which could be used as one dimension in an assessment centre process.

In addition the exercises can be used as part of a development process for staff, or during career or outplacement counselling. The initial score provides an indication of the level of performance of the individual. Working through the scenarios and dilemmas in greater detail with a coach or mentor would provide a wealth of opportunities to discuss the individual's work style and how his or her approach might be improved.

The tests can also be used to evaluate training. For example, staff who are new to a customer service role could take the customer service SJT prior to the training and then again following the course in order to ascertain the effectiveness of the training and learning.

Chapter 2: Using IRIS

Scoring and Norms

The items within IRIS tests are scored by comparing them to 'ideal response' as agreed by subject matter experts within the relevant field. The feedback report gives an overall percentile score, allowing comparisons between candidates to be drawn, and summarises the candidate's 'Response Style', such as 'Supporting Others' and 'Results Focus'. This style can provide further detail on what the candidate prioritises when responding to a situation.

A number of different norms have been developed for the customer service and graduate tests including norms based on job incumbents and norms based on a representative sample of the general population who have no experience in the role. Please contact the TalentLens team for more details of these.

New norms can be created for use with the IRIS tests. Please contact the TalentLens team to find out more about this service.

Preparing candidates for testing

Candidates should be informed why they are being asked to take the test, the overall purpose of the test in the selection process and how they can access the test. This should include a description of the test and access to example items. These materials should be sufficient to allow candidates with disabilities to determine whether they are able to access the exercise effectively or need some form of accommodation.

Candidates should also be provided with information about how their results will be used and what feedback they will receive on their results.

Where candidates will be taking the test unsupervised it is important to encourage them to respond honestly and on their own. This can be achieved by explaining the benefit to them as well as the organisation of a good skills match with the job. In addition candidates should be reminded of any steps in the selection process to verify scores. This can include a feedback discussion on the results where candidates may be required to explain their approach and additional assessments such as role plays or interview questions which will reassess the same areas in more depth.

Chapter 3: Results and Interpretation

Following completion of the IRIS test online, an interpretation report will be generated which will support the user in interpreting the results. These are appropriate for sharing with line managers and other personnel who are not test trained.

Using scores

A common use of the IRIS tests is in short-listing candidates to be invited to a more in-depth assessment process. The ability of the IRIS tests to provide a general indication of the appropriateness of the candidate's judgement in dealing with typical situations that arise at work makes them an excellent choice for this purpose. Short-listing can be achieved by selecting candidates who score above a given level or selecting the top scoring candidates according to the desired number to be invited to the next stage of assessment. Both provide a straightforward method of reducing candidate numbers and ensuring that more suitable candidates are seen at the next round.

Whichever approach is used it is important that a single instrument is not given too much weight in any selection decision. This is because a single factor will only explain a part of work performance and by being too selective on one factor, it will be more difficult to select on another factor. For instance being too selective using IRIS scores might mean that there will be little choice in terms of analytic ability or work experience.

Some users may prefer to use the IRIS tests in a more qualitative manner. In this case the score on the IRIS tests is viewed together with other information about the candidate to decide how well the candidate meets the job requirements. This approach is more appropriate when there are relatively small numbers of candidates or at the later stages of selection.

For development purposes scores may be helpful in showing the suitability of a candidate for a different role. For instance a version of the IRIS tests designed for a sales or customer service role might be helpful in deciding whether someone working in an administrative role might have potential for working in a more customer facing environment.

Where an individual has performance difficulties, the dilemmas within the IRIS exercises may provide useful material for discussion to help the person understand how their performance might be improved. The use of the IRIS exercises may encourage a less emotionally charged

discussion of appropriate behaviour than confronting the individual with mistakes they have made in their work.

Test scores should be interpreted carefully. Scores can be affected by a test taker's state, for example, anxiety or feeling unwell. Candidates with a disability or with English as a second language may be disadvantaged due to the test format. For these reasons scores should be explored carefully and interpreted with caution. On occasion, test scores may contradict alternative information on a test taker. In this case, the test user should work with the test taker to explore the information and discover possible causes for these anomalies.

The IRIS tests have been carefully standardised using standard administration procedures. Any changes to this process can result in unreliable test scores. Used correctly, psychometric tests are a powerful tool that can provide important information on the test taker but for these reasons tests are designed to be used alongside alternative assessment methods. It would not be appropriate to use scores on IRIS exercises for purposes such as selecting people for redundancy or performance appraisal. Assessment of performance for these purposes should be based on more direct information.

Where there are score differences between groups, very high cut scores can result in adverse impact against lower scoring groups. (See chapter 4 for discussion of group scoring patterns for IRIS tests).

Users should be aware of the accuracy of IRIS scores and take account of the typical band of error around scores before making over fine distinctions between levels of performance.

Where candidates have completed the tests without supervision, it is important to take steps to verify the scores at later stages of assessment. This means that exercises which reassess the same areas of competence should be included. Making it clear to candidates that scores are verified through other exercises will tend to discourage candidates from engaging in cheating behaviours such as asking someone else to complete the test for them.

Interpretation

The candidate's score is an indicator of their suitability for the role when the IRIS version has been well matched to job requirements. It is important to remember that the IRIS tests only measure some aspects of job aptitude. They assess the candidate's judgement of appropriate behaviour in response to a variety of situations. They do not measure the candidate's reasoning skills. While the scenarios behind the IRIS questions reflect real life dilemmas for employees, the test only measures whether candidates differentiate appropriate from inappropriate responses. They do not directly measure whether candidates would respond appropriately in real life. It is important to use other assessment methods to verify the suitability of candidates in addition to IRIS. It is never good practice to rely on a single assessment tool to make decisions about applicants for jobs.

By analysing the responses chosen, it is possible to explore the individual's style of responding when completing IRIS. This '**Response Style**' is used in the customer service and graduate tests to interpret a respondent's performance (see examples of reports for these tests on the next two pages). The response style indicates particular patterns of responses and is independent of the individual's overall performance on IRIS. For example, if an individual's style is to look for agreement with others, this may mean that more responses are chosen when completing IRIS that reflect seeking agreement, consensus or compromise. In reality it is likely that an interaction between the person's style, their ability, and the nature of a work situation will determine how the person responds to a situation. Response styles are norm-referenced, but should only be used for feedback and developmental guidance as each scale is only based on the selection of responses which relate to it.

Taking into account the above caveats, IRIS test scores can be a good indicator of future performance as shown by the validation studies described in the next chapter.

Candidates with low scores are unlikely to perform well in the role. Their judgement and behaviour choices are likely to deviate from the desirable on a frequent basis. They are likely to need a high degree of supervision and training to develop and maintain appropriate responses. They are likely to work at a low level for a number of the competency areas described.

Candidates with moderate scores are likely to perform at an adequate level in a well matched role. Their judgement of what is appropriate behaviour in different situations is likely to be generally sound although they may err on occasion. They will require typical levels of supervision and training to maintain and develop appropriate work behaviours.

While there may be areas of competency in which they do not meet the required standard, for the most part their competency levels will be adequate or above.

Candidates with high scores have shown a good understanding of appropriate behaviour for a well matched role. They are likely to show good judgement in dealing with a variety of situations and will make fewer poor decisions than most, although like everyone, they are likely to make errors from time to time. They are likely to be able to work well with minimal levels of supervision and will be able to benefit from training to develop their skills further.

Customer Service IRIS Feedback Report – 2 page sample

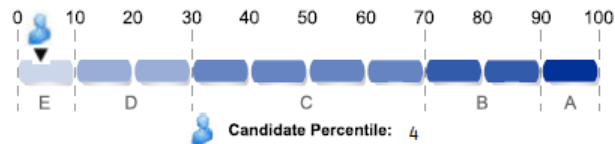


Background

IRIS is a situational judgement test which assesses judgement when faced with commonly occurring work situations. In completing IRIS you may recall that you completed three tasks with a range of work-related situations. The situations included issues relating to solving problems, working within guidelines and procedures, supporting a team and interacting with customers. You were asked to rate the appropriateness of a number of responses to these situations. Your IRIS score reflects your ability to use effective judgement to resolve these and similar complex work situations.

Overall Performance - Your score is as follows

Your responses were scored and your result was compared with that of other people working in customer service representative roles. In comparison to this group, your performance on IRIS suggests your awareness to judge the appropriateness of responses to situations is Well below average. Your percentile score means you scored equal to or greater than 4% of people working in customer service representative roles.



- A - Well above average** (91st percentile and above)
- B - Above average** (71st – 90th percentiles)
- C - Average** (31st – 70th percentiles)
- D - Below average** (11th – 30th percentiles)
- E - Well below average** (10th percentile and below)

Suggestions of things you might do if you would like to improve your performance

- Think about how you have dealt with similar situations in the past. What worked well? What could you improve upon?
- Talk to colleagues about how you might deal with challenging situations. Analyse these situations and share ideas.
- Ask other people you have worked with for feedback regarding your approach. What do they think you do well? What could you do better?
- Think about someone you know who deals with work situations well. Try to analyse their approach. What could you learn from them?

name A Sample page 2

Response Style

Throughout IRIS you were asked to evaluate the appropriateness of responses to a number of work situations. The appropriateness of responses usually depends on the context of the situation.

By analysing the responses you chose, it is possible to explore your style of responding when working through IRIS. It is important to point out that your style indicates particular patterns of responses only -and that this is independent of your overall performance on IRIS as reported on page 2.

For example, if a person's style or preference is to look for agreement with others, this may mean that more responses are chosen when completing IRIS that reflect seeking agreement, consensus or compromise. In reality it is likely that an interaction between the person's style, their ability, and the nature of a work situation will determine how the person responds to a situation. In this example, style is but one of a number of factors contributing to their overall judgement or performance.

This information is not used as part of the assessment process, but rather is intended for feedback purposes only.

Response styles are reported in comparison to a large sample of people who have completed IRIS.

Your response style summary

- Focus on quality:** Your response style suggests you may place moderate emphasis on quality when dealing with situations, compared to others.
- Supporting others:** Your response style suggests that, compared to others, you may place less emphasis on supporting others in situations.
- Assertiveness:** Your response style suggests you may place more emphasis on being assertive compared to others.

name A Sample page 3

Graduate IRIS Feedback Report – 2 page sample

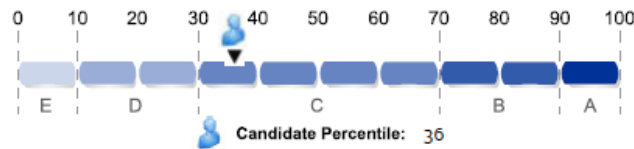


Background

IRIS is a situational judgement test which assesses judgement when faced with commonly occurring work situations. In completing IRIS you may recall that you completed three tasks with a range of work-related situations. The situations included issues relating to managing people, solving problems, working within guidelines and procedures, and responding to change. You were asked to rate the appropriateness of a number of responses to these situations. Your IRIS score reflects your ability to use effective judgement to resolve these and similar complex work situations.

Overall Performance -Your score is as follows

Your responses were scored and your result was compared with that of other UK Graduates and Junior Managers. In comparison to this group, your performance on IRIS suggests your awareness to judge the appropriateness of responses to situations is Average. Your percentile score means you scored equal to or greater than 36% of UK Graduates and Junior Managers.



- A - Well above average (91st percentile and above)
- B - Above average (71st – 90th percentiles)
- C - Average (31st – 70th percentiles)
- D - Below average (11th – 30th percentiles)
- E - Well below average (10th percentile and below)

Suggestions of things you might do if you would like to improve your performance

- Think about how you have dealt with similar situations in the past. What worked well? What could you improve upon?
- Talk to colleagues about how you might deal with challenging situations. Analyse these situations and share ideas.
- Ask other people you have worked with for feedback regarding your approach. What do they think you do well? What could you do better?
- Think about someone you know who deals with work situations well. Try to analyse their approach. What could you learn from them?

name A Sample page 2

Response Style

Throughout IRIS you were asked to evaluate the appropriateness of responses to a number of work situations. The appropriateness of responses usually depends on the context of the situation.

By analysing the responses you chose, it is possible to explore your style of responding when working through IRIS. It is important to point out that your style indicates particular patterns of responses only - and that this is independent of your overall performance on IRIS as reported on page 2.

For example, if a person's style or preference is to look for agreement with others, this may mean that more responses are chosen when completing IRIS that reflect seeking agreement, consensus or compromise. In reality it is likely that an interaction between the person's style, their ability, and the nature of a work situation will determine how the person responds to a situation. In this example, style is but one of a number of factors contributing to their overall judgement or performance.

This information is not used as part of the assessment process, but rather is intended for feedback purposes only.

Response styles are reported in comparison to a large sample of people who have completed IRIS.

Your response style summary

- Consulting:** Compared to others, your response style suggests you may place **moderate** emphasis on consulting with others when decision making or problem solving.
- Sensitivity:** Your response style suggests that compared to others, you may place **moderate** emphasis on being sensitive to the feelings of others when problem solving.
- Toughness:** Your response style suggests you may place **moderate** emphasis on being tough on poor behaviour, compared to others.
- Results focus:** Your response style suggests you may place **moderate** emphasis on achieving results when problem solving, compared to others.

name A Sample page 3

Sales IRIS Feedback Report – 2 page sample

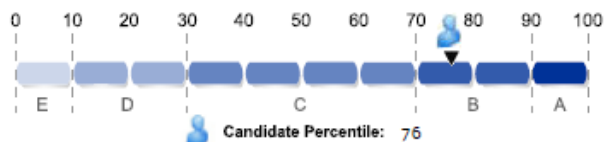


Background

IRIS is a situational judgement test which assesses judgement when faced with commonly occurring work situations. In the IRIS test candidates are asked to complete eight tasks across a range of work-related situations. The situations included issues relating to managing new enquiries, advancing the sale process, navigating gatekeepers, understanding the customer, prospecting and time management. Respondents are asked to rate the appropriateness of a number of responses to these situations. The IRIS score below reflects Ann Example's ability to use effective judgement to resolve these and similar complex work situations.

Overall Performance

Ann Example's score has been compared with Sales Professionals who have previously taken the test. In comparison to this group, Ann Example's performance on IRIS suggests their ability to judge the appropriateness of responses to situations is above average. The percentile score below means Ann Example's score is equal to or greater than 76% of Sales Professionals.



- A - Well above average (91st percentile and above)
- B - Above average (71st – 90th percentiles)
- C - Average (51st – 70th percentiles)
- D - Below average (31st – 50th percentiles)
- E - Well below average (10th percentile and below)

Suggestions of ways Ann Example might improve performance

- Think about how you have dealt with similar situations in the past. What worked well? What could you improve upon?
- Talk to colleagues about how you might deal with challenging situations. Analyse these situations and share ideas.
- Ask other people you have worked with for feedback regarding your approach. What do they think you do well? What could you do better?
- Think about someone you know who deals with work situations well. Try to analyse their approach. What could you learn from them?

name Ann Example

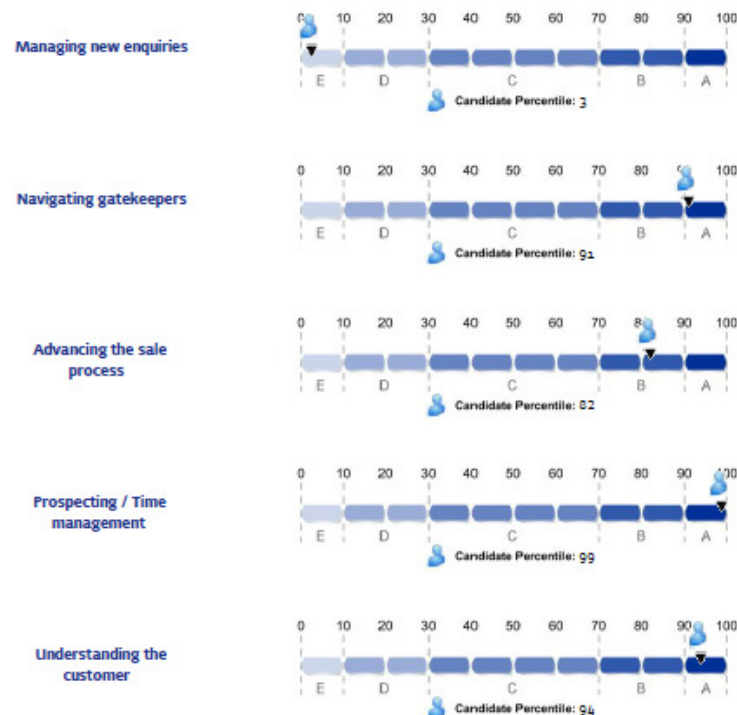
page 2

Task Performance

Throughout IRIS, Ann Example was asked to evaluate the appropriateness of responses to a number of typical sales related activities. Below is a summary of Ann Example's performance across a number of different situations that reflect a particular theme or type of task.

Further to the overall score, Ann Example's results in these areas were compared with Sales Professionals who have previously taken the test in order to give a percentile score for each area. The percentile scores for these sub-areas should not be interpreted in isolation, but give an indication of performance on the test across a number of different areas. A low score on a particular subscale may indicate a potential area for development.

Ann Example -Task performance summary



name Ann Example

page 3

Chapter 4: Technical Information

Development

The development procedure for the IRIS tests followed a rigorous process reflecting best practice. A team of experienced test developers worked together to create an innovative measure which captures the best the method allows. As the IRIS tests are a generic measure it was important to identify situations which epitomised work dilemmas common across many roles. To this end, subject matter experts (SMEs) were consulted widely during the development process, both individually and in focus groups. These were individuals who trained or supervised people in the target roles. They provided ideas for scenarios and responses and were used in determining the keyed responses. Where there was no clear consensus on the scoring between the experts, questions were dropped.

Following the initial market research, the content for the three tests was designed based on the SMEs ideas for potential scenarios and responses. This work was carried out by external consultants with prior experience of situational judgement tests. This design process included creating the scenarios, the possible responses to the scenarios and the initial scoring key.

Following this, the items and scoring keys were tested with SMEs. This allowed for the examination of the most frequently chosen responses to a situation. Focus groups of SMEs were used to follow up in the cases where the SMEs disagreed with the test developers on the responses to a situation. This process allowed for the refining of the scoring key. Both public and private sector participants were sampled for the SMEs and Focus Group stages.

The pool of selected items was extensively trialled on job incumbents from both internal Pearson employees and external clients from both the public and private sector. Items were analysed to evaluate the difficulty level of each item, how well each item differentiated between better and worse respondents, score differences between minority and majority groups and correlations with job performance. Based on the outcome of this analysis, final items were selected for inclusion in the tests.

Scoring

There are a number of different approaches to scoring situational judgement tests. This depends to some extent of the format of the test itself. The format of the IRIS tests allows candidates to review multiple responses to a situation. Each response is rated on a 4 point scale. Candidates receive more points the closer their response is to the ideal response for the questions.

Ideal responses were developed with the aid of subject matter experts. These are individuals who are very familiar with the target job families. Typically they are supervisors or managers of individuals in relevant roles. Only items where there was a consensus between different subject matter experts and between the subject matter experts and test developers were included in the final tests.

To assess the agreement regarding the ideal response to items, concordance data was collected from samples of the target roles within a number of organisations. Item-level analysis was conducted to ensure that there was a high level of agreement in the final items included in each test.

Reliability

Reliability is an indication of the accuracy of measurement. One of the advantages of using properly developed psychometric measures is that they have known levels of accuracy. For bespoke assessments, the level of accuracy is rarely known.

The single most used measure of reliability is the internal consistency reliability which provides a lower bound estimate of the likely accuracy of scores from a test with different content. Values above 0.7 are desirable for measures to be used in short-listing. The internal consistency reliability of the Customer Service IRIS test was found to be 0.76.

Validity

The validity of a test is the extent to which it measures what it is intended to measure.

Content Validity

As the test items have been developed with job incumbents and subject-matter experts in the relevant area, the content validity of the IRIS tests is high. Any items that did not reflect the role in question were removed from the test during the early stages of development.

Face Validity

The IRIS SJT tests have high face validity as it is clear to respondents that they are measuring their potential to perform effectively in the role in question. 95% of experienced sales professionals who completed the IRIS Sales SJT felt that the scenarios accurately reflected those typically undertaken in a sales role, with 98% reporting that the test is relevant to a sales role.

Criterion Validity

A number of concurrent validation studies have been carried out which have shown that the IRIS tests are effective in predicting job performance with correlations ranging from 0.23 to 0.33 across a range of performance criteria. One such study with a customer service team found that the Customer Service test has a correlation of 0.27 with ratings of overall job performance, and the 'Assertiveness' response style has a correlation of 0.24 with managerial ratings of assertiveness.

The correlation of the IRIS tests with other instruments was examined. There were weak correlations with various ability tests showing the IRIS test scores are only mildly influenced by the intelligence of respondents. The test showed weak correlations with some personality scales. This reflects the broad measurement aims of the questions.

Group Comparisons

One of the advantages of the situational judgement approach is that there are generally small if any differences between candidates from different gender, ethnic or age backgrounds. The IRIS tests were developed to be accessible to people of all backgrounds with and without previous job experience. The tests have been trialled on diverse groups including members of different ethnic groups, of either gender and varying age and sexual orientation. Where samples were sufficient the results were broken down by group membership and the results compared.

Overall there were few differences between groups and where these did occur they tended to be of small to moderate size.

References

McDaniel, M.A., Morgeson, F.P., Finnegan, E.B., Campion, M.A. & Braverman, E.P. (2001). Predicting job performance using situational judgement tests: A clarification of the literature. *Journal of Applied Psychology, 86*, 730-740.

Motowidlo, S.J. and Van Scotter, J.R. (1994). Evidence that task performance should be distinguished from contextual performance. *Journal of Applied Psychology, 79*(4), 475-480.

Robertson, I. T. and Smith, M. (2001). Personnel Selection. *Journal of Occupational and Organizational Psychology, 74* (4), 441–472.