

REDEFINE EMPLOYEE SUCCESS

Exousia HR

People Performance Measurement and Management System

Creating successful employees and teams presents a challenge for any company.

Exousia HR delivers accurate and reliable estimation about the ability of each employee.

Exousia creates quantitative Success Profiles for each job or role, identifies gaps, and suggests ways how to improve employees' knowledge –

thus helping employees to be successful in their occupation and assisting companies in building and maintaining high performing teams.

Exousia leverages innovative mathematical approaches automatically and at scale.

This paper provides an overview of Exousia capabilities and presents a [real-life case study](#).

*Exousia comes from the Greek word **ἐξουσία** (eck-so-see-a), that means “power” and “authority.”*



Contents

Introduction.....	3
The Difficulty of Items (Questions)	3
Rating Scales.....	3
Foundational Items	3
Exousia – Overview and Benefits	4
CASE STUDY	5
Agreement among Raters	5
Ability of Employees, Difficulty of Items	6
Difficulty of Items	6
Ability of Employees	7
Item Characteristic Curve	8
Causal Relationships among Items	9
Employee Success Profile	10
Employee Proficiency Cards	10
Comparison of Employees.....	12
Team Strengths and Gaps	13
Strengths and Gaps in the Team	14
Personalized Programs for Employees.....	14
Conclusion.....	15

© 2019 YieldWise Inc. This document is for informational purposes only.

YIELDWISE MAKES NO WARRANTIES, EXPRESS OR IMPLIED, IN THIS SUMMARY.

The names of actual companies and products mentioned herein may be the trademarks of their respective owners.

Introduction

A Success Profile for identifying high performing individuals – either within an organization, existing teams or for future hires – outlines two critical parts of individuals to be successful in their jobs and to be aligned with the business future needs:

- Capability (technical knowledge)
- Accountability (responsibilities, problem-solving, etc.)

Success Profile development process is based on knowledge assessments. The result of the analysis of these assessments is a Success Profile that describes differentiating knowledge and skills of successful performers of a specific job. Unfortunately, knowledge assessments are often incorrectly analyzed, and it leads to wrong inferences about people ability regarding a particular role.

The reasons for incorrect analysis are often originating from the misuse of assessments of raw scores. When people are assessed through a series of tests, it is tempting to manipulate raw test data with simple mathematics immediately. However, researchers agree that the use of raw scores to evaluate and compare employees' achievements is erroneous¹.

The Difficulty of Items (Questions)

Do all items (questions) in an assessment require the same amount of efforts from people to respond to them? The answer is “No,” as it is highly unlikely that all items are of similar difficulty. Therefore, the use of sum or average of raw scores will be misleading in the evaluation of the respondent fit to the job. In some cases, items are assigned points (weights) reflecting the different difficulty of items. However, the magnitude of difficulty of assessment items is not an absolute value, and it depends on the ability of people taking the test. Measuring employees' ability, while disregarding actual items difficulty, creates misleading results. To address this problem, Exousia uses Polytomous Rasch Measurement Model which correctly analyzes raw test scores by simultaneously estimating the difficulty of items and ability of respondents.

¹ Wright BD, Stone MH. Best Test Design. Chicago: MESA Press; 1979

Rating Scales

Assessment tests commonly use rating scales, such as Strongly Agree (SA), Agree (A), Disagree (D), and Strongly Disagree (SD). The responses usually coded as 4, 3, 2, and 1 (e.g., SA is a 4, A is a 3), and it can be very appealing to conduct mathematical operations with those numbers. One problem with performing mathematical operations with numerically coded rating-scale answers is an assumption that the interval from a Strongly Agree to Agree is the same as from Agree to Disagree. It is true that Strongly Agree denotes more agreement than Agree, and that Agree means more agreement than Disagree, and so on. However, this does not immediately mean that the intervals between adjacent categories are equal.

Item 1	SA	A	D	SD
Item 2	SA	A	D	SD
Item 3	SA	A	D	SD

In addition to differences of intervals between adjacent rating categories, the pattern of the intervals is usually different from item to item. Such data is called “ordinal” data. Moreover, all items should not be assumed to be equally agreeable. For example, a 3 (Agree) in response to item 1 should not be assumed to indicate the same level of agreement as answering a 3 (Agree) to item 2.

The problem with the analysis of rating scales is addressed by Exousia through the application of Polytomous Rasch Measurement Model. Specifically, it uses test respondents' raw scores and expresses the respondents' performance on a linear scale that accounts for the unequal difficulties across all test items so that that true person ability can be computed.

Foundational Items

In any assessment test, knowledge of different topics is interdependent. We cannot assume that the knowledge required to answer one item is independent of the knowledge needed to answer other items.

Let's say one question in the test validates knowledge of a specific technique required for the job, and two other questions validate usage of that technique to solve problems. The deficiency of knowledge in the first question leads to failure in the two other questions. In contrary, sufficient knowledge in the first question enhances chances of success in the two related questions – thus the first question is considered to be foundational. Such relations among test items (questions) are not always straightforward and obvious; they may include dependencies on more than one item and are thus not easily detectable. Identifying foundational items allows detecting areas of knowledge and experience essential for success in a specific role. Exousia uses Relational Bayesian Networks to address this problem successfully.

Exousia – Overview and Benefits

Exousia methodology, methods, algorithms and software solutions were developed to extract actionable insights from psychometric and knowledge test data. Exousia can be accessed to analyze assessment data through a simple and straightforward API or just through a data interface.

Exousia provides six core functionalities that are not available in traditional methods of employees' performance measurement and management:

- **Agreement Among Raters**

In the evaluation of employees' performance usually involved several raters -- the employees themselves, their managers, peers, etc. The relationship between raters, for example, the employee and the manager, has a significant impact on employee success in the organization. Professional or personal disagreement may lead to significant problems for the team and even for the organization as the whole. Early identification of such disagreement may help to resolve the issue before it becomes a problem. Exousia uses Cohen Kappa statistics to identify employees that disagree with their manager, as well as identify items that cause this disagreement.

- **The Ability of Respondents and the Difficulty of Items**

Exousia's modified Polytomous Rasch Measurement Model (PRMM) can process

incomplete data (e.g., missing values) and provide a robust estimation of items' difficulty and respondents' ability. Exousia enables managers to evaluate the right fit of the employees to their role. At the same time, Exousia helps improve the quality of knowledge tests by identifying and eliminating malfunctioning items.

- **Causal Relationships Among Items**

Exousia can identify causal relationships among knowledge test items and reveal which items are foundational. This functionality is implemented through Relational Bayesian Networks (RBN) methodology and proprietary algorithms that build RBN structure from the responses to knowledge test. Identification of foundational items and their relationship with other items, as well as with respondents' ability, contribute to the development of Success Profile and creation of personalized improvement plans for employees.

- **Success Profile**

Using the results of PRMM and RBN, the proprietary algorithm of Exousia creates Success Profiles for each job role. The most important outcome of it is the quantitative assessment of the qualitative attributes of a Success Profile. Using Success Profiles, managers and recruiters will be able to identify employees' and candidates' ability threshold to ensure their success in the job.

- **Employee Proficiency Cards**

Exousia automatically builds Employee Proficiency Cards that contain person proficiency per each assessment test item and serve as a basis for determination of the employee's overall fit to the job. Exousia automatically creates recommendations for personalized plans to improve employees' ability, to close existing gaps, and to increase chances of successful performance in the role.

- **Teams Strengths and Gaps**

Exousia aggregates data from individual Employee Proficiency Cards to evaluate the team's competency regarding job performance. Managers and HR representatives use this information help the team in achieving current and future goals in the organization.

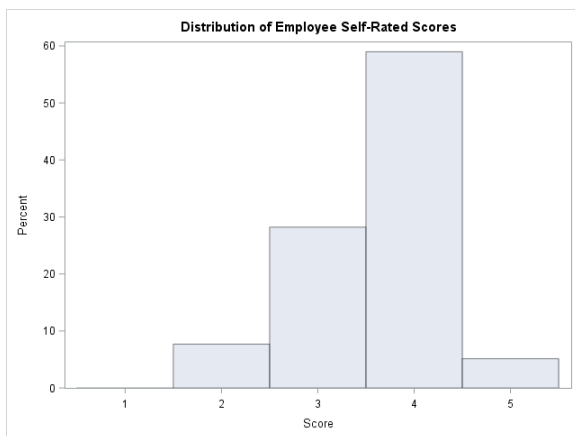
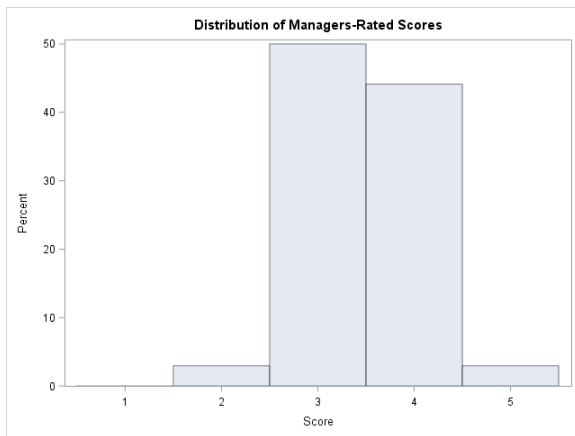
CASE STUDY

This case study was conducted on data obtained during a quarterly performance review of 33 employees from data scientists' teams. Employees were asked to rate themselves on a 5-level scale with regards to their knowledge and experience in 23 knowledge areas required to perform their job. Managers were also asked to rate the employees on the same items.

The five levels were coded as numbers in the following way:

- Some Awareness (no knowledge) - 1
- Novice (limited experience) - 2
- Intermediate (practical application) - 3
- Advanced (applied theory) - 4
- Expert (recognized authority) - 5

The HR department performed the analysis of employee self-rated scores and manager-rated scores obtained from this performance review. They simply calculated averages of scores per employee and received results presented as histograms below.



Although fundamentally incorrect, this method is still used quite often. Averaging the scores ignored the ordinal nature of the data, the different difficulties of the items, the disagreement of employees and managers on scores for particular items. The outcome outlined differences in the score but didn't reveal the sources of those differences. No recommendations on how to improve team performance and proficiency were made.

The company decided to use Exousia to analyze the scores to knowledge assessment test -- employees self-assigned (rated) scores and managers rated scores.

Agreement among Raters

As the first step, Exousia identifies which items caused disagreement between employees and managers. These items are sources of discrepancy and should be excluded from further analysis. Cohen Kappa statistics identified that 9 (gray shaded) out of 23 items caused disagreement between the employees and the managers. These nine items were removed from further analysis.

Table 1. Items in Agreement/Disagreement

#	Item	Agreement with Manager
1	Visualization 1	In Disagreement
2	Statistical Analysis 2	In Agreement
3	Data Gap Identification 3	In Agreement
4	Analytics Workflow 4	In Disagreement
5	Programming & Coding 5	In Agreement
6	Data Monetization 6	In Agreement
7	Programming & Coding 7	In Agreement
8	Business Intelligence 8	In Disagreement
9	Data Monetization 9	In Agreement
10	Feature Engineering 10	In Agreement
11	Business Intelligence 11	In Agreement
12	Theoretical Quantitative Foundations 12	In Agreement
13	Analytics Workflow 13	In Disagreement
14	Data Analysis & Interpretation 14	In Disagreement
15	Data Analysis & Interpretation 15	In Agreement
16	Statistical Analysis 16	In Disagreement
17	Visualization 17	In Disagreement
18	Business Intelligence 18	In Disagreement
19	Model Validation 19	In Agreement
20	Feature Engineering 20	In Agreement
21	Munging data 21	In Disagreement
22	Model Validation 22	In Agreement
23	Tech Savvy 23	In Agreement

Based on the remaining 14 items, Exousia identified which employees are in disagreement with their managers about the rating scores. 16 out of 33 employees, appear to be in disagreement. For 16 employees in disagreement with managers, Exousia performed the analysis based on both self-rated and manager-rated scores. For the 17 employees who are in agreement with the manager, Exousia used managers-rated scores for the analysis. In this paper, we present results only for the employees that are in agreement with managers according to Cohen Kappa statistical criterion.

Table 2. Employees Rating in Agreement/Disagreement

#	Employee	Agreement with Manager	#	Employee	Agreement with Manager
1	E10335	In Disagreement	18	E21013	In Disagreement
2	E10336	In Disagreement	19	E21014	In Agreement
3	E10337	In Disagreement	20	E21015	In Agreement
4	E10338	In Agreement	21	E21016	In Agreement
5	E10343	In Disagreement	22	E21017	In Disagreement
6	E10344	In Disagreement	23	E21018	In Disagreement
7	E10345	In Agreement	24	E21019	In Agreement
8	E10347	In Agreement	25	E21020	In Disagreement
9	E10348	In Agreement	26	E22867	In Agreement
10	E10349	In Agreement	27	E22868	In Disagreement
11	E10350	In Agreement	28	E23595	In Disagreement
12	E21007	In Agreement	29	E26674	In Disagreement
13	E21008	In Disagreement	30	E26675	In Disagreement
14	E21009	In Agreement	31	E35903	In Agreement
15	E21010	In Agreement	32	E75121	In Agreement
16	E21011	In Agreement	33	E75122	In Disagreement
17	E21012	In Disagreement			

Ability of Employees, Difficulty of Items

It is the nature of the assessment framework that questions (items) of the test have *different difficulty*, and the respondents have *different ability* regarding the test items. Exousia estimates the difficulty of test questions (items) that reflect the level of proficiency of the employee in the item. At the same time, Exousia assesses each employee's ability relative to the test. The polytomous Rasch Measurement Model (PRMM) estimates difficulty and ability *simultaneously*. As a result, items are ranked according to their difficulty, and the employees are ordered according to their ability.

Difficulty of Items

The difficulty of assessment items, estimated by PRMM, reflects the level of complexity of the item for the rated team. In our case study, 17 employees were rated by the managers. Items with lower

difficulty generally do not present a challenge for the team, and items with higher difficulty appear to be arduous to the employees.

The PRMM not only estimates the difficulty of the items but also identifies OutFit value, an outlier-sensitive fit. An OutFit value greater than 1.3 indicates that the difficulty level of the item does not relate consistently to the ability of employees. For example, if generally high-rated employees were rated very differently on the item (some rated high while some were rated low) this indicates inconsistency. The similar discrepancy may appear when out of the mostly lower-rated employees some received high scores and some – low scores for the same item.

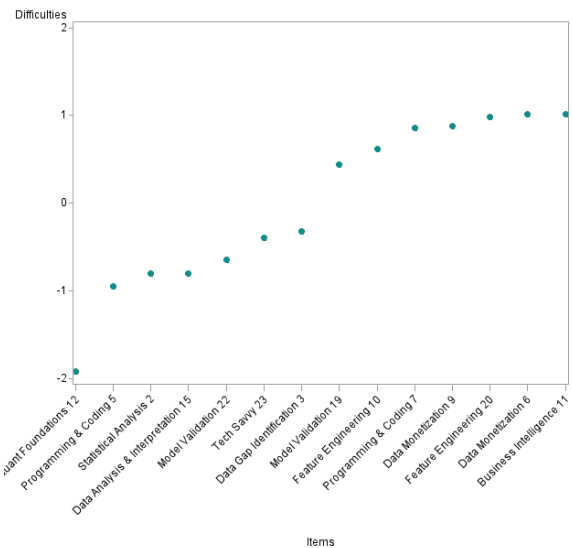


Table 3. Items Difficulty

#	Item	Difficulty	OutFit
1	Theoretical Quant Foundations 12	-1.92	1.47
2	Programming & Coding 5	-0.95	0.70
3	Statistical Analysis 2	-0.80	0.67
4	Data Analysis & Interpretation 15	-0.80	0.64
5	Model Validation 22	-0.64	1.30
6	Tech Savvy 23	-0.39	1.72
7	Data Gap Identification 3	-0.32	0.99
8	Model Validation 19	0.44	0.91
9	Feature Engineering 10	0.62	0.66
10	Programming & Coding 7	0.86	2.08
11	Data Monetization 9	0.88	0.29
12	Feature Engineering 20	0.99	0.67
13	Data Monetization 6	1.01	0.75
14	Business Intelligence 11	1.01	1.04

There are three items shaded gray in Table 3 for which the OutFit value greater than 1.3, thus they appear to be malfunctioning items – it is possible that the questions were not formulated clearly, may

contain unfamiliar terms, or may have other reasons leading to misunderstanding by respondents. Exousia suggests excluding these items from the evaluation of employees' proficiency.

Ability of Employees

While making conclusions about the employee's performance, it is critical to estimate their ability with consideration of the difficulty of test questions (items). Such a conclusion reflects *the real proficiency* of the employee, in opposite to raw scores.

The employees' ability is estimated by PRMM conditionally on items difficulty: lower numbers mean lower employee ability relatively to the assessment, and higher numbers indicate higher ability.

An OutFit value higher than 1.3 indicates employees for whom some of the observed scores are too far from the expected values estimated by PRMM. It suggests that, in some items, the employee's proficiency was possibly over- or underestimated by the managers. In Table 4, there are three employees (shaded gray) for whom OutFit is greater than 1.3.

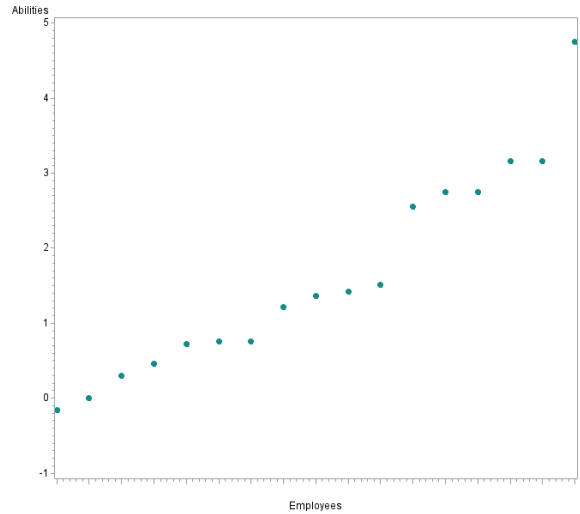
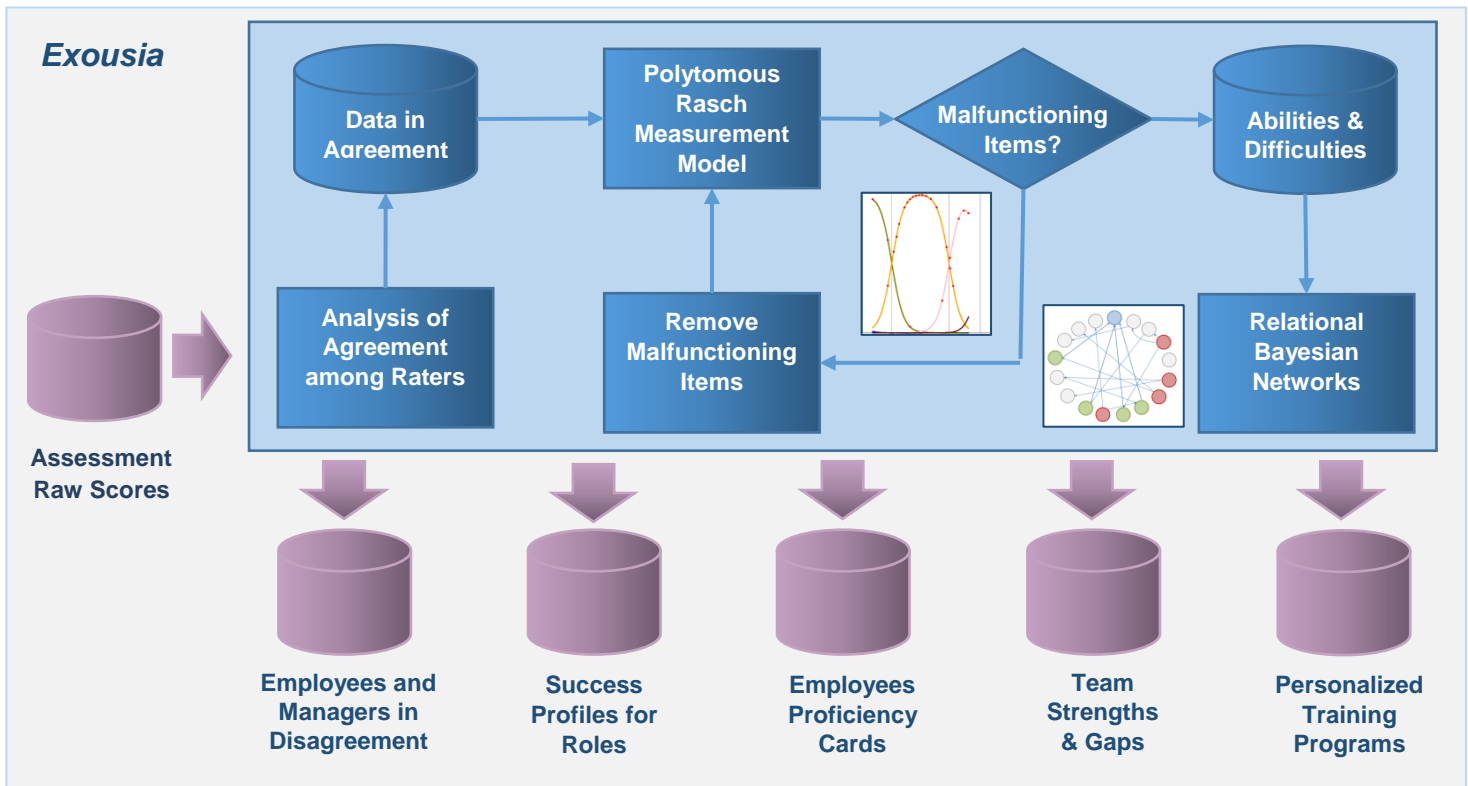


Table 4. Employees Ability

#	Employee	Ability	OutFit	#	Employee	Ability	OutFit
1	E10349	-0.15	1.29	10	E21014	1.42	0.45
2	E21007	0.00	0.751	11	E35903	1.52	1.26
3	E21016	0.31	1.40	12	E75121	2.56	0.90
4	E21009	0.46	0.38	13	E10348	2.75	1.15
5	E21015	0.73	0.70	14	E21019	2.75	0.28
6	E10345	0.76	1.27	15	E10338	3.16	0.33
7	E22867	0.76	2.33	16	E10347	3.16	1.87
8	E21011	1.21	0.75	17	E10350	4.75	0.53
9	E21010	1.37	1.03				



Item Characteristic Curve

Item Characteristic Curves (ICC) describe the relationship between the ability of employees and the estimated probability of the employees to get a specific score. ICCs are created by PRMM. Each item in the assessment has its ICC, and for each item, Exousia identifies the probabilities of each employee to receive a particular score.

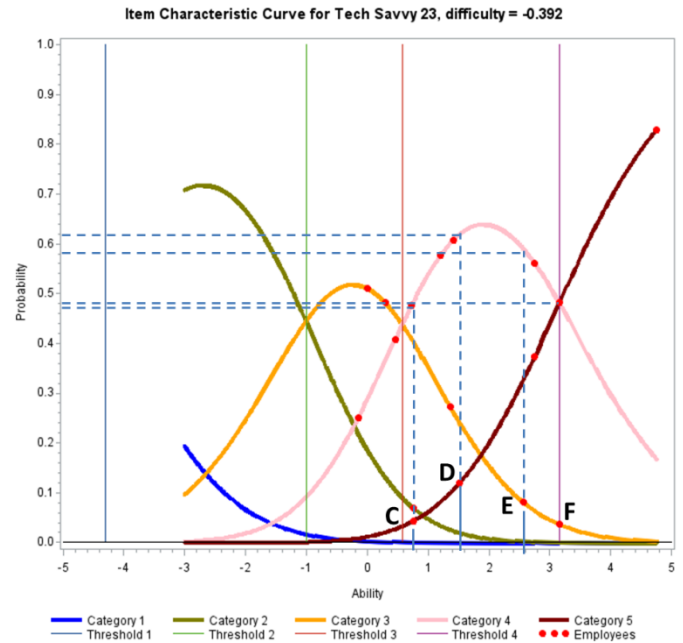
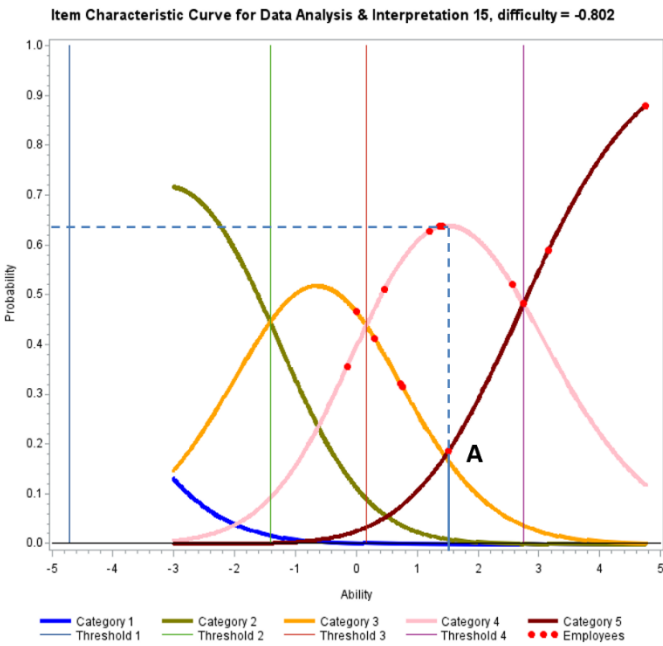
For example, the following ICC is built for Data Analysis & Interpretation (item 15). Each curve of the ICC represents a probability for an employee to get a specific score (Category), depending on their ability. Thresholds (solid vertical lines) identify ability for which the probabilities of adjacent scores are equal. For example, the pink curve represents a probability distribution to get score 4 for this item. According to Threshold 3 (red solid vertical line), when an employee has ability = 0.14, they have equal chances to receive score 3 (yellow curve) or 4. Thus, to get score 4, an employee should have the ability higher than 0.14. Red dots on the curves denote actual employees' scores.

Thus, the employee was rated higher than their actual proficiency on item 15.

For Tech Savvy (item 23), presented on the following graph, ICC tells a different story. We already know that this item was identified as malfunctioning due to a high OutFit value. ICC below exposes some details of Exousia decision to mark this item as malfunctioning. Points C, D, E and F on the ICC illustrate employees who have a high probability of getting different scores than they received:

Point	Employee	Ability	Actual Score/ Probability	Expected Score / Probability
C	E22867	0.76	5 / P = 0.04	4 / P = 0.48
D	E35903	1.52	5 / P = 0.12	4 / P = 0.62
E	E75121	2.56	3 / P = 0.08	4 / P = 0.59
F	E10347	3.16	3 / P = 0.04	4 or 5 / P = 0.48

Exousia builds ICCs for all items and estimates what level of employees' ability can ensure professional success in the role.



Point A is located on the burgundy curve (Category 5) and represents employee E35903 with ability = 1.52. This employee was rated 5, while for their ability the probability of getting this score is only 0.19. The broken blue vertical line reaches the purple curve (Category 4) in the point where the corresponding probability of getting score 4 is 0.64.

Causal Relationships among Items

The understanding of causal relations among items enables the identification of what *foundational knowledge* drives success in the assessment. Exousia uses Relational Bayesian Networks (RBN) analysis to identify probabilistic *causal relationships* among test items and employees' ability. RBN visualizes dependency or influence of one item on another as a graph. The arrows in the graph pointing from one item to another reflect how knowledge in one item impacts competency in another item.

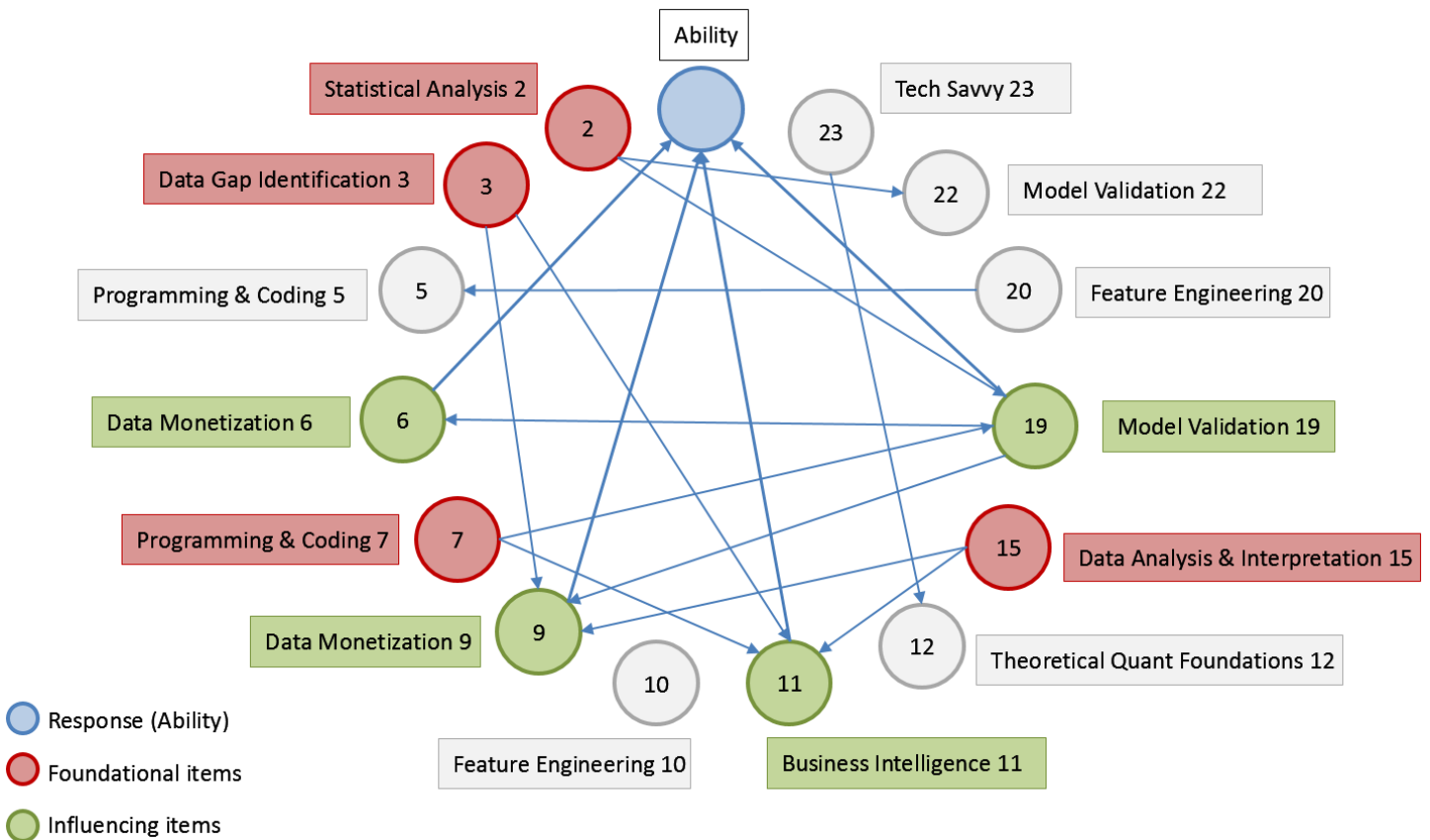
In this case study, according to the RBN built by Exousia, the employees' ability directly influenced (thick blue arrows originated from green colored items) by their knowledge in four items:

- Data Monetization 6
- Data Monetization 9
- Business Intelligence 11
- Model Validation 19

However, for the employees' performance evaluation, it is important to identify which items are foundational.

Exousia identifies foundational items that are critical for the improvement of employees' proficiency in their job. The following four items are identified as *foundational*:

- Statistical Analysis 2
- Data Gap Identification 3
- Programming & Coding 7
- Data Analysis & Interpretation 15



Employee Success Profile

Using the results produced by PRMM and RBN, Exousia builds Employee Success Profile for the data scientist role. Employee Success Profile determines which scores employees should receive for each item to be successful in their job. The Success Profile states:

- which items should be considered for success, and which should be excluded (gray shaded)

items were identified as malfunctioning and were eliminated from the Success Profile)

- what is the lowest score an employee should receive for each item to be proficient, and
- which of the items tested during the assessment are foundational (red bordered) for an employee’s success in their job.

Employee Success Profile for Data Scientist Role

#	Item	Difficulty	Score	Item Importance
1	Theoretical Quant Foundations 12	-1.92	5	
2	Programming & Coding 5	-0.95	5	
3	Data Analysis & Interpretation 15	-0.80	5	Foundational
4	Statistical Analysis 2	-0.80	5	Foundational
5	Model Validation 22	-0.64	5	
6	Tech Savvy 23	-0.39	4	
7	Data Gap Identification 3	-0.319	4	Foundational
8	Model Validation 19	0.44	4	
9	Feature Engineering 10	0.62	4	
10	Programming & Coding 7	0.86	4	Foundational
11	Data Monetization 9	0.88	4	
12	Feature Engineering 20	0.99	4	
13	Business Intelligence 11	1.01	4	
14	Data Monetization 6	1.01	4	

Employee Proficiency Cards

In this case study, the approach based on average scores estimated that almost 50% of employees received score 3, and about 45% of employees received score 4. What is a sufficient score to be successful in this job? Are all employees with score 4 have the same proficiency in their role? Are those employees on the path of success or failure? Exousia provides answers to these questions as it automatically builds Proficiency Cards for each employee based on estimated items’ difficulty and employees’ ability, as well as on the course Success Profile and ICCs.

The Proficiency Cards identify an employee’s level of competency in each item, and relate it to the importance of the item and the required level of knowledge in each item:

- **Strength** – the employee exceeds the score requirement for the item in the Success Profile,
- **Fit** – the employee is in line with the score requirement for the item in the Success Profile,
- **Opportunity to Fit** – the employee has a high probability of fitting the score requirement of the Success Profile,
- **Gap** – the employee’s actual and expected scores are below the score of the Success Profile.

Employee Proficiency Cards allow identifying differences between actual scores and expected scores estimated by PRMM. This helps to identify “hidden” gaps or “hidden” opportunities – for example, cases when the employee might have been mistakenly assigned the higher score that was not in line with their ability, or the employee was assessed lower score than it was estimated based on their ability.

Employee Proficiency Cards serve as a basis for the determination of every *employee proficiency in the job*:

- **Exceeds Proficiency (Expert)** – the employee exhibits Strength in all items,
- **Proficient (Advanced)** – the employee shows Strength or Fit in all items,
- **Foundational Proficiency (Intermediate)** – the employee shows Strength or Fit in all foundational items of the Success Profile,
- **Partially Proficient (Novice)** – the employee shows Strength, Fit or Opportunity to Fit in all foundational items,
- **Insufficient Proficiency (Some Awareness)** – the employee didn't classify to any of the four groups above.

Employee Proficiency Cards contain the following data:

- Gray-shaded items are excluded from consideration as they were identified as malfunctioning.

- Red-bordered items were identified as foundational.
- Light-green cells identify the highest probability of score for each item.
- The column “Actual Score” contains the score the employee received on the specific item.
- The column “Most Likely Score” contains the score that is most probable for the employee according to PRMM (see the probabilities in the light-green shaded cell).

Examples below demonstrate Proficiency Cards for two employees.

The employee E10338 with ability = 3.16 is *Partially Proficient* as they exhibit Fit or Opportunity to Fit in all foundational items according to the Success Profile. Exousia revealed that the employee was scored lower than required by the Success Profile on one of the foundational items: Statistical Analysis 2.

Proficiency Card for Employee E10338, Ability 3.16, Partially Proficient

Item	Actual Score	Prob. Receiving Score 1	Prob. Receiving Score 2	Prob. Receiving Score 3	Prob. Receiving Score 4	Prob. Receiving Score 5	Most Likely Score	Success Profile	Status
Business Intelligence 11	4	0.00	0.01	0.20	0.64	0.16	4	4	Fit
Data Analysis & Interpretation 15	5	0.00	0.00	0.02	0.39	0.59	5	5	Fit
Data Gap Identification 3	4	0.00	0.00	0.04	0.50	0.46	4	4	Fit
Data Monetization 6	4	0.00	0.01	0.20	0.64	0.16	4	4	Fit
Data Monetization 9	4	0.00	0.01	0.17	0.64	0.18	4	4	Fit
Feature Engineering 10	4	0.00	0.01	0.13	0.63	0.23	4	4	Fit
Feature Engineering 20	4	0.00	0.01	0.19	0.64	0.16	4	4	Fit
Model Validation 19	4	0.00	0.00	0.11	0.62	0.27	4	4	Fit
Model Validation 22	5	0.00	0.00	0.03	0.43	0.55	5	5	Fit
Programming & Coding 5	5	0.00	0.00	0.02	0.36	0.63	5	5	Fit
Programming & Coding 7	4	0.00	0.01	0.17	0.64	0.18	4	4	Fit
Statistical Analysis 2	4	0.00	0.00	0.02	0.39	0.59	5	5	Opportunity
Tech Savvy 23	4	0.00	0.00	0.04	0.48	0.48	4	4	Fit
Theoretical Quant Foundations 12	5	0.00	0.00	0.00	0.18	0.82	5	5	Fit

Another employee, E10348, with ability = 2.75 demonstrates *Foundational Proficiency* as they show Strength or Fit in all foundational items of the Success Profile.

Proficiency Card for Employee E10348, Ability 2.75, Foundational Proficiency

Item	Actual Score	Prob. Receiving Score 1	Prob. Receiving Score 2	Prob. Receiving Score 3	Prob. Receiving Score 4	Prob. Receiving Score 5	Most Likely Score	Success Profile	Status
Business Intelligence 11	3	0.00	0.03	0.28	0.60	0.10	4	4	Opportunity
Data Analysis & Interpretation 15	5	0.00	0.00	0.04	0.48	0.48	4 or 5	5	Fit
Data Gap Identification 3	4	0.00	0.00	0.07	0.57	0.35	4	4	Fit
Data Monetization 6	3	0.00	0.03	0.28	0.60	0.10	4	4	Opportunity
Data Monetization 9	4	0.00	0.02	0.25	0.62	0.11	4	4	Fit
Feature Engineering 10	3	0.00	0.01	0.20	0.64	0.15	4	4	Opportunity
Feature Engineering 20	4	0.00	0.03	0.27	0.60	0.10	4	4	Fit
Model Validation 19	3	0.00	0.01	0.17	0.64	0.18	4	4	Opportunity
Model Validation 22	4	0.00	0.00	0.05	0.52	0.44	4	5	Gap
Programming & Coding 5	5	0.00	0.00	0.03	0.45	0.52	5	5	Fit
Programming & Coding 7	5	0.00	0.02	0.25	0.62	0.12	4	4	Strength
Statistical Analysis 2	5	0.00	0.00	0.04	0.49	0.48	4 or 5	5	Fit
Tech Savvy 23	5	0.00	0.00	0.06	0.56	0.37	4	4	Strength
Theoretical Quant Foundations 12	5	0.00	0.00	0.01	0.25	0.75	5	5	Fit

Comparison of Employees

Exousia can differentiate employees that are impossible to distinguish using averaging of scores. Two presented above Proficiency Cards scored the employees differently.

Table 5. Comparison of Employees

#	Items	Employee E10338	Employee E10348	Success Profile
1.	Business Intelligence 11	4	3	4
2.	Data Analysis & Interpretation 15	5	5	5
3.	Data Gap Identification 3	4	4	4
4.	Data Monetization 6	4	3	4
5.	Data Monetization 9	4	4	4
6.	Feature Engineering 10	4	3	4
7.	Feature Engineering 20	4	4	4
8.	Model Validation 19	4	3	4
9.	Model Validation 22	5	4	5
10.	Programming & Coding 5	5	5	5
11.	Programming & Coding 7	4	5	4
12.	Statistical Analysis 2	4	5	5
13.	Tech Savvy 23	4	5	4
14.	Theoretical Quant Foundations 12	5	5	5
	AVERAGE SCORE	4.29	4.14	
	EXOUSIA SCORE	Partially Proficient	Foundational Proficiency	

employees E10338 and E10348 using these two methods. The results of this comparison are in Table 5.

The calculation of average would assign the employee E10338 score of 4.29, and the employee E10348 – score 4.14. These scores are both high and very close one to another. The inference would be that both employees are similarly proficient in their jobs.

However, Exousia revealed that these employees are significantly different. Table 5 demonstrates that the employee E10348 is in Fit or Strength with all the foundational items (cells shaded blue), and thus exhibits Foundational Proficiency in their role. The employee E10338 is in fit with only three out of four foundational items (cells shaded orange), and therefore only Partially Proficient. Exousia identified gaps in foundational knowledge of the employee E10338 (yellow cell), which allowed creating a personalized training plan for that employee.

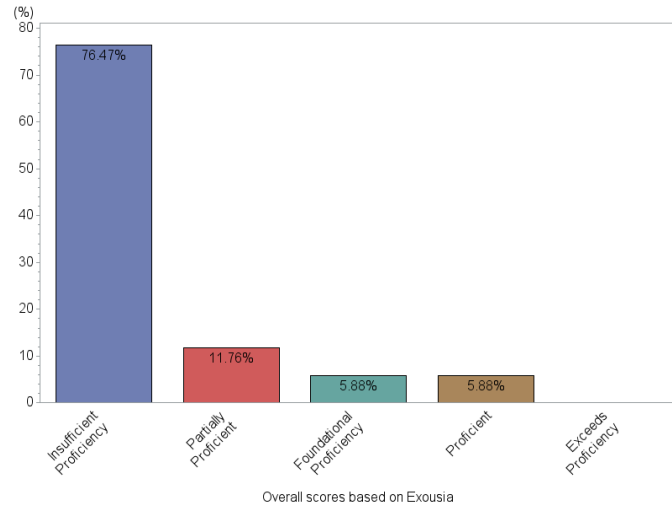
The approach based on averaging of scores *failed in identifying this critical difference* and missed the fact that the employee E10338 has a significant gap in the foundational knowledge area.

To demonstrate the superiority of Exousia over averaging of scores, let's compare evaluation of

Team Strengths and Gaps

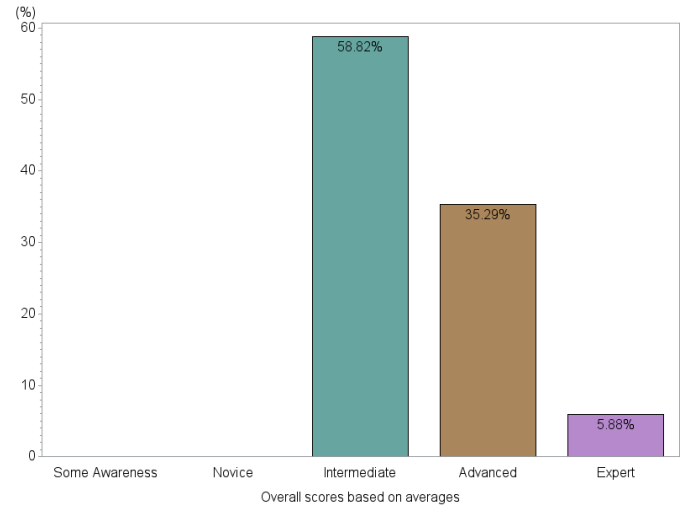
Success Profile serves as a basis for employees' proficiency assessment. It determines scores an employee needs to achieve to be successful in their job. It identifies foundational items and these items serve as a differentiator in the assessment process.

- 5.88% of employees are Proficient,
- 5.88% of employees demonstrated Foundational Proficiency,
- 11.76% of employees are Partially Proficient,
- 76.47% of employees have Insufficient Proficiency



Proficiency scores calculated by Exousia vary significantly from the scores based on the averaging approach.

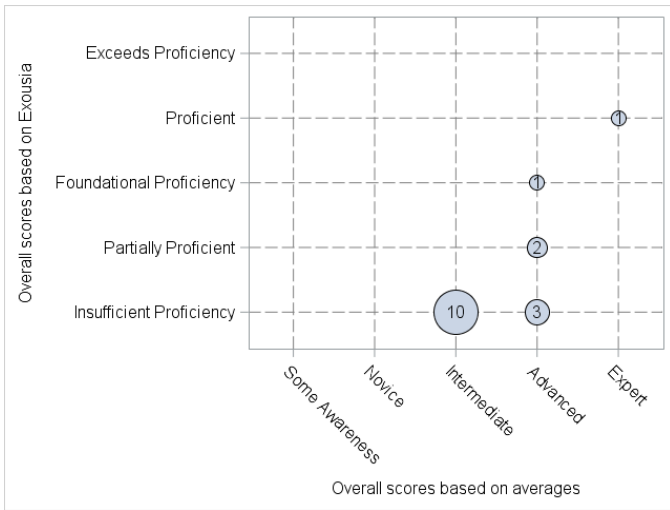
According to the averaging method, none of employees received scores lower than 3 (Intermediate). Such results convey a reassuring message, essentially stating that almost all employees are of the average or higher proficiency.



Exousia discovered a very different situation. It identified foundational items – the knowledge areas that are most important for the success in the role. It also identified the lowest score per each item that an employee needs to achieve to be successful in the job. Exousia revealed that 13 out of 17 employees (76.47%) demonstrated gaps in at least one foundational item and thus have insufficient proficiency in the required knowledge areas. These employees are set up to fail in their role. Exousia identified knowledge gaps, created recommendations for training programs, and thus provided an opportunity for the employees to improve their skills. The averaging method *failed to identify* employees with insufficient proficiency.

The comparison of overall average scores vs. Exousia scores for each employee (see the figure below) revealed the following:

- Ten employees, scored as Intermediate by averaging method, received Insufficient Proficient score by Exousia as they exhibited Gap in at least one foundational item.
- One employee scored as Advanced demonstrated Foundational Proficiency.
- Two employees scored as Advanced were Partially Proficient.
- Three employees scored Advanced exhibited Insufficient Proficiency.
- The employee scored Expert appeared to be Proficient by Exousia.

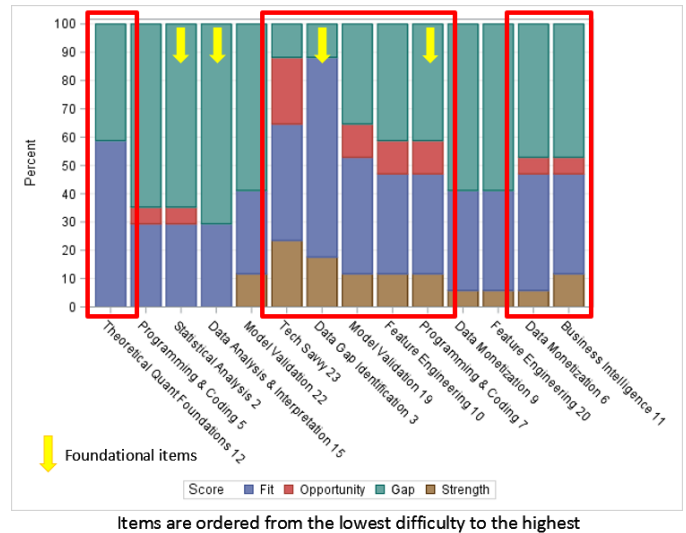


- Foundational items identify the most critical areas of knowledge, and in two of them, Statistical Analysis 2 and Data Analysis & Interpretation 15, less than 30% of employees exhibited Fit.
- Eight items framed in red outline knowledge areas in which at least 50% of employees exhibit Strength, Fit or Opportunity to Fit.
- There are six areas of knowledge where the team needs to improve their skills.

In the area of individual scoring, the essential advantage of Exousia is its ability to discover hidden gaps in foundational knowledge areas. This helps in identifying personalizing learning programs for employees to ensure their professional success.

Strengths and Gaps in the Team

Exousia identifies areas of overall strengths and gaps of the team, thus providing actionable feedback to the manager. The bar chart visualizes the following insights:



Personalized Programs for Employees

Exousia automatically creates recommendation of customized programs to improve employees' ability, to close existing gaps, and to increase chances of success in the job. For each foundational item identified by Exousia, the employees for whom reinforcement is required (red) or optional (green) were identified.

Personalized Training Program for the Data Scientists Team

Employee	Data Gap Identification 3	Programming & Coding 7	Statistical Analysis 2	Data Analysis & Interpretation 15
E10349	Required	Required	Required	Required
E21009	Required	Required	Required	Required
E10345		Required	Required	Required
E21007		Required	Required	Required
E21010		Required	Required	Required
E21011		Required	Required	Required
E21014			Required	Required
E21015			Required	Required
E21016			Required	Required
E21019			Required	Required
E22867			Required	Required
E75121		Optional		Required
E35903		Required		
E10338			Optional	
E10347		Optional		

Conclusion

Commonly used averaging of scores to evaluate employees' proficiency fail short of expectation. In this case study, we showed that this approach misses important gaps in employees' skills and produces misleading results. Exousia offers the solution that provides an accurate assessment of employees' proficiency along with actionable insights on how to improve employees' skills to ensure their success in the job.

Exousia allows to:

1. Identify which employees disagree with their managers regarding the rating scores so that reasons for the disagreement will be reviewed and the employees will receive a fair performance evaluation.
2. Rank employees by their *ability conditionally on the difficulty* of the assessment items.
3. Identify the *quality of the assessment structure* (in terms of items), as well as which topics are *foundational* for the employees' success in the job.
4. Measure *overall employees' proficiency* in the knowledge areas assessed by the test.
5. Create the *Employee Success Profile* for each job/role and determine the realistic scores an employee should receive for each item to become successful in his career.
6. Create *Employee Proficiency Cards* which report the level of *proficiency for each employee per each item* as well as *overall competence* of the employee regarding the performance assessment.
7. Create a *training plan for each employee*.

According to the commonly used average measurement of proficiency, employees get scores that allow them to pass the assessment and move on. However, the reality could be different, and the employees could have hidden gaps in knowledge areas essential for their professional success. Exousia helps to reveal the real situation with employees score in the assessment. Exousia provides correct and accurate measurement of employees' success and creates actionable recommendations on how to close revealed gaps. This helps the companies to develop better and more successful teams, identify and treat performance before they become problems, and eventually dramatically reduce employees' attrition.