SILESIAN UNIVERSITY OF TECHNOLOGY PUBLISHING HOUSE

SCIENTIFIC PAPERS OF SILESIAN UNIVERSITY OF TECHNOLOGY 2021
ORGANIZATION AND MANAGEMENT SERIES NO. 153

THE IDENTIFICATION OF PROBLEMATIC SOFTWARE TESTERS MANAGEMENT AREAS

Alina LANGA[1*], Ireneusz J. JÓŹWIAK[2], Kacper STAROŚCIAK[3], Jan SWITANA[4], Alicja M. JÓŹWIAK[5], Wojciech M. NOWAK[6]

[1] Wroclaw University of Science and Technology, Faculty of Computer Science and Management, Wrocław; alina.langa2@gmail.com
[2] Wroclaw University of Science and Technology, Faculty of Computer Science and Management, Wrocław; ireneusz.jozwiak@pwr.edu.pl, ORCID: 0000-0002-2160-7077
[3] Wroclaw University of Science and Technology, Faculty of Electronics, Wrocław; kacper.starosciak@gmail.com
[4] Wroclaw University of Science and Technology, Faculty of Microsystem Electronics and Photonics, Wrocław; switana.jv@mail.com, ORCID: 0000-0001-7188-8199
[5] Wroclaw University of Science and Technology, Faculty of Architecture, Wrocław; ajozwiak07@gmail.com, ORCID: 0000-0002-8796-2549
[6] Wroclaw University of Science and Technology, Faculty of Architecture, Wrocław; wojciech98nowak@gmail.com, ORCID: 0000-0003-4861-8566
* Correspondence author

Purpose: The purpose of this paper is to explore the condition of managing manual software testers.
Design/methodology/approach: The analysis is based on a survey targeting three groups (managers, testers, development team members) that will touch on areas consistent with the management functions.
Findings: The condition of testers' management can be described as good, but there is a probability of improvement for the areas that have been studied.
Research limitations/implications: It's worth carrying out similar research on larger samples to make sure about the obtained averages values and statistical differences between them. It's also worth proposing methods to improve the possible problematic areas that have been found in this research paper.
Practical implications: This paper may draw the attention of managers whose subordinates are software testers that there are areas where they can improve their management.
Originality/value This research shows how testers' management is perceived by managers, how it is perceived by testers and how it is perceived by other members of development teams.

Keywords: software tester management, management function, manual testing.

Category of the paper: Research paper.

1. Introduction

Software testing is often treated dismissively. It's not only a separate opinion of the authors of this article but it's also pointed out in the literature on the subject. For example, it's noted (Roman, 2015) that while it can be said that software testing exists as long as its development, many organizations only pay attention to employing high-class programmers, at the side of which, unfortunately, testers with both low experience and commitment still happens to work, because their competences and qualifications are not so comprehensively verified. Even testers working in such a huge and significant organization as Google weren't treated equally, until they have learned the techniques of tests automation (Whittaker et al., 2012). Managers often forget that this approach leads to high risk of releasing applications with critical defects, which consequently gives rise to significant additional costs.

The key core of software testing management is software testers management. This is caused by the specificity of their work, because software testing is hard to observe. He's not a producing worker – he's a worker who controls what others have produced. Tester shouldn't therefore be assessed directly for the effects that are visible from the outside, but for his general attitude to the position as well as approach to the duties entrusted to him. It's related to the differences between human psychology and testing (Certyfikowany tester, ISTQB, 2018) because informing about errors detected by the tester may be perceived by other people involved in the product as a reprimand, although it should be a natural part of working on software. The measurability of the effects is often in conflict with diplomacy in communication. Tester should not show satisfaction with the errors he found, to avoid offending the programmers. On the other hand, for managers, sometimes it's the only determinant of whether the tester has done his job. If there are no bugs in the application, the tester's work may go unnoticed.

The research carried out at the initial stage of this work didn't reveal the existence of any research concerning the diagnosis of the condition of software testing management. It was decided to identify areas of testers management that make problems.

2. Management functions in the context of testers management

Management functions are logical groups of activities related to each other. Continuous performance of them is one of the basic duties of every manager (Koźmiński, Piotrowski, 2007). These functions are planning, organizing, leading and controlling (Griffin, 2004). Different authors distinguish management functions differently, for example by adding human resources (Koontz, O'Donnell, 1972) to them or identification (Stabryła, 2018). Others omit leading,

limiting them to planning, organizing and controlling Steinmann, Schreyogg, 2001). Some authors expand this classification even further, listing as many as nine functions: planning, organizing, information and knowledge management, organization financial management, operations management, personnel policy, marketing and public relations, negotiations, controlling (Jemielniak, Koźmiński, 2011). The first of the typologies, that is planning, organizing, leading and controlling, was considered by the authors of this work to be the most universal and adequate to the problem discussed in it, therefore it will be subject to further considerations.

Planning can be defined as setting goals and ways to achieve them (Griffin, 2004). It allows gathering tips on how to act. It's a kind of starting point for all activities undertaken. It doesn't only mean setting goals, but also the ways in which they will be monitored and controlled. The main difference between monitoring planning and controlling planning is that monitoring planning determines how progress will be checked, and controlling planning determines how the effect of this work will be checked. With regard to the management of testers, it seems interesting to figure out what action should be taken by the tester and what should be achieved by these actions, because his work is often highly creative. The set goals should be measurable, specific and unambiguous by definition, so there is a supposition that such formulation of goals may be difficult when managing testers. It may also be problematic to figure out ways to monitor and control work, also because of their frequent non-normative character.

Organizing is a logical grouping of activities and resources (Griffin, 2004). It consists in coordinating and sharing work – so as not to create conflicts and properly allocate human resources. It should allow the manager to be relieved of the duties that have been entrusted to him, as long as his employees have the appropriate competences to be able to fulfill them. It is also extremely important to create an appropriate communication network between employees so that they are as well informed as possible about what co-workers are doing, which prevents duplication of work and gives a picture of the current situation of activities. The quality of the testers work organization can be recognized by comparing the planned time for the tests with the actual duration time. It can also be recognized that the division of labor between the testers is right – if no one is overloaded, if no one is obliged to do many small tasks and simultaneously focus their attention on different disjoint areas of action.

Leading, that is the most important and most ambitious managerial activities. They are used to encourage employees to cooperate so that it benefits their organization (Griffin, 2004). The manager should, through his interpersonal skills, ensure good communication with his subordinates. He should also motivate in such a way as to have a positive impact on employees for both sides. In the case of testers management it is particularly important to give consent to going beyond schemes – while guarding that these schemes were also filled. It is worth paying attention to the existence of an interesting form of leadership, which is situational leadership (Blanchard et al., 2019). It consists in the appropriate adaptation of the management style depending on the competences of the employee and the commitment he puts in the work.

There is a suspicion that the situational leadership is rarely used to manage the testers, which can be invaluable loss for many organizations.

Controlling is watching the progress of the objectives (Griffin, 2004). Here there is a direct reference to the planning, the implementation of the established monitoring and control there. With regard to tester management, it can be considered whether the work of testers is actually measurable, because only by meeting this condition can we talk about the proper implementation of this function.

In summary, each of these management functions can be easily applied in theory to the management of testers. It may be more difficult, however, to perform them in reality. It should be remembered that in order for the use of management functions to bring measurable results, it's necessary to use them all at the same time, skillfully intertwining them.

3. Research goal and design of the survey

The research goal is to identify problematic areas of testers management. These areas will be divided according to the management functions, as well as more detail, narrowing them to specific managerial challenges. To achieve this will be used one-time survey.

The research material will be three groups of thirty people professionally related to software testing. The first group included in the research will be managers whose subordinates are testers. The next group will be just software testers. The last group consists of people cooperating with testers, so they will be other members of project teams.

Summing up, 90 people will be surveyed. The estimated return of surveys sent via LinkedIn, a social networking site that specializes in professional and business contacts, is about 5% (Stokes et al., 2019). This means that to get 90 completed questionnaires, should contact about 1,800 people. Despite such a low level of feedback, it was decided to distribute the survey by LinkedIn because it is considered a reliable source of information about the professional experience of users (Paliszkiewicz, 2018), which will allow the survey to be directed to the appropriate target group. Attempts will be made to increase the feedback rate by designing a questionnaire with a short completion time, that would be about 5 minutes, and by personalizing invitations to complete the survey sent in private messages. The survey will be anonymous with an optional possibility to sign.

The questionnaire will contain a list of closed questions. The answer scale used for all questions will be a five-point Likert scale (Kaczmarek, Tarka, 2013): 1) definitely not, 2) probably not, 3) I don't know, 4) probably yes, 5) definitely yes.

The list of questions, divided into categories, is as follows:

1. Planning

P1 – Is in your team defined what actions the tester should take?

P2 – Is in your team defined what the tester should achieve with his actions?

P3 – Is in your team defined how the tester's activities will be monitored – how will the progress of work be checked?

P4 – Is in your team defined a way in which the tester's activities will be controlled – how will the results of the work be checked?

2. Organizing

P5 – Is there equality in your team between estimation of the time needed for tests and how much time actually needs to be spent on them?

P6 – Is the division of labor between the testers in your team appropriate?

P7 – Does it happen in your team that testers interfere with each other, for example if they are working on the same test environment?

P8 – Does the tester in your team has the appropriate competences, some managerial responsibilities and decision-making powers delegated to him by the manager?

P9 – Is communication between testers in your team good?

P10 – Is communication between testers and other team members in your team good?

3. Leading

P11 – How good is communication between testers and the manager in your team?

P12 – Is in your team openness to ideas and creativity of testers?

P13 – Is there a different treatment due to competences and commitment of testers in your team?

P14 – Is there a proper motivation for testers to do their work?

4. Controlling

P15 – Is tester's work measurable in your team?

P16 – Are tester's activities monitored in your team (is the progress of tester's work always checked)?

P17 – Are in your team tester's activities controlled (is the result of tester's work always checked)?

4. Analysis of the results

Collecting responses to the survey was launched on 6 October 2020. A higher than expected feedback, around 20%, allowed to complete the collection of responses on 27 October 2020.

The survey was attended by 30 managers, 30 manual testers and 30 other team members. In the last mentioned group were programmers, business analysts, UX and UI designers, system architects and consultants. Managers and team members work directly with manual software testers.

Valuing answers for the P7 question was reversed for purposes of analysis because it has the opposite sense than all the other questions.

The overall average of all responses without the division into target groups and categories was 3,71, which means that the overall assessment of tester management quality is the closest to the answer "Probably yes". Therefore, it is assessed more positively than negatively. The variance was 1,45. We can define the variance as quite large, so the responses are varied. The obtained skewness is -0,83. This indicates a left-hand distribution, which can be interpreted so that most of the respondents placed the assessment above average.

Average, variance and skewness for the results of the research divided into categories are presented in the table 1. The highest average was achieved for the planning area and the lowest for controlling. At the same time, the highest variance was obtained for controlling and the lowest for leading. Also recorded the highest skewness for leading, and the lowest for organizing.

Table 1.
Average, variance and skewness of the results divided into categories

	Average	Variance	Skewness
Planning	3,80	1,50	-0,90
Organizing	3,76	1,44	-0,91
Leading	3,64	1,34	-0,65
Controlling	3,60	1,56	-0,81

Source: own work.

It would be interesting to know whether the variation in the results obtained in these categories varies significantly. To determine this, a one-way variance analysis ANOVA will be performed using the Excel Data Analysis tool at the significance level $\alpha = 0,05$. The test hypotheses are as follows:

H0: $\mu_{Planning} = \mu_{Organizing} = \mu_{Leading} = \mu_{Controlling}$

H1: it's not true that H0

The resulting table 2 provides a summary – the test statistics such as the number of elements, sums, averages, and variances. In the table 3 a variance analysis is presented, including the sum of squares SS, degrees of freedom df, square averages, F value, p-value and a critical F value for a given df. The value that we are primarily interested in is p-value. It's higher than the significance level, that is, we don't have the basis to reject the zero hypothesis. On the basis of the sample, we can therefore conclude that, at the accepted significance level, the differences between the averages in the populations of the categories under consideration are not statistically significant.

Table 2.
One-way ANOVA for categories – summary

Groups	Counter	Sum	Average	Variance
Planning	360	1368	3,8	1,497493
Organizing	540	2028	3,755556	1,435498
Leading	360	1312	3,644444	1,34398
Controlling	270	973	3,603704	1,563555

Source: own work.

Table 3.
One-way ANOVA for categories – variance analysis

Sources of variance	SS	df	MS	F	p-value	F test
Between groups	8,620044	3	2,873348	1,980081	0,115033	2,610734
Within groups	2214,419	1526	1,451126			
Total	2223,039	1529				

Source: own work.

The table 4 presents averages, variances and skewness for the survey results divided into a target groups. The highest average and response skewness were achieved among managers and the lowest among testers. Conversely was the response variance – the lowest was among managers and the highest was among testers.

Let's check whether the differences between the averages in the population surveyed by target groups are statistically significant. We will also use a one-way ANOVA at the significance level $\alpha = 0,05$. The test hypotheses are as follows:

H0: $\mu_{Managers} = \mu_{Testers} = \mu_{OtherTeamMembers}$

H1: it's not true that H0

Table 4.
Average, variance and skewness of the results divided into target groups

	Average	Variance	Skewness
Managers	4,09	1,15	-1,37
Testers	3,44	1,57	-0,52
Other team members	3,61	1,42	-0,74

Source: own work.

The resulting summary is shown in the table 5 and the variance analysis is shown in the table 6. The p-value was set to 0 by the tool[1], so is very small much less than the significance level. To make sure, it has also been calculated with another, more accurate tool and amounted to 1,1102e-16. We can therefore reject a zero hypothesis in favor of an alternative hypothesis and conclude that the differences between the averages in populations of the samples are statistically significant.

[1] One-way ANOVA post-hoc Tukey HSD Calculator, https://astatsa.com/OneWay_Anova_with_TukeyHSD/.

Table 5.
One-way ANOVA for target groups – summary

Groups	Counter	Sum	Average	Variance
Managers	510	2086	4,090196	1,147055
Testers	510	1753	3,437255	1,574641
Other team members	510	1842	3,611765	1,416757

Source: own work.

Table 6.
One-way ANOVA for target groups – variance analysis

Sources of variance	SS	df	MS	F	p-value	F test
Between groups	116,566	2	58,28301	42,24985	0	3,001617
Within groups	2106,473	1527	1,379484			
Total	2223,039	1529				

Source: own work.

The additional post-hoc Tukey HSD analysis, carried out using the same tool used to ensure the p-value, showed that there are statistically significant differences in the averages between each of the three pairs of groups, namely between managers and testers (H0: $\mu_{Managers} = \mu_{Testers}$, H1: it's not true that H0, p-value = 0,0010053), between managers and other team members (H0: $\mu_{Managers} = \mu_{OtherTeamMembers}$, H1: it's not true that H0, p-value = 0,0010053) and even between testers and other team members (H0: $\mu_{Testers} = \mu_{OtherTeamMembers}$, H1: it's not true that H0, p-value = 0,0467403).

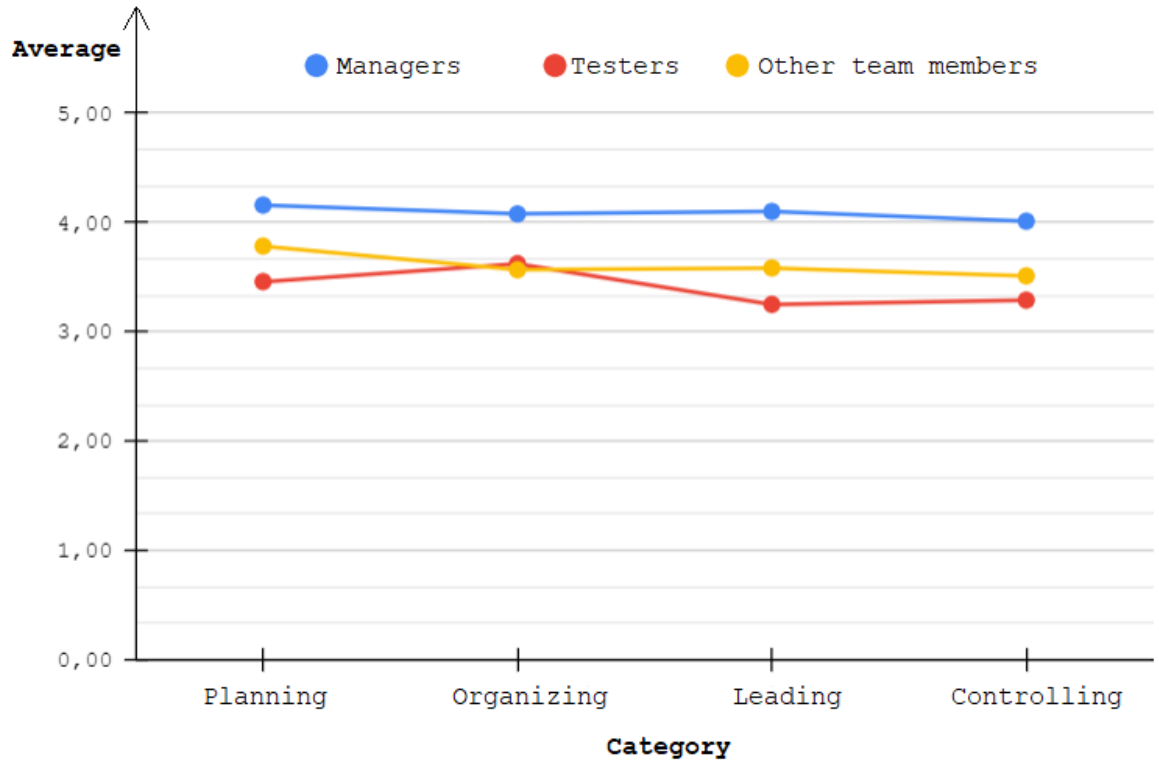

Figure 1. The chart of average results for categories by target groups. Source: own work.

As we start the analysis at a lower level and now have more detailed results to compare, so to make the differences more visible, a chart (fig. 1) of the average results for the categories by group is prepared based on the data in the table 7. The chart shows that managers gave the average highest responses in each category. The lowest average responses were provided by testers, except for organizing where the other members gave the lowest responses.

Table 7.
Average, variance and skewness of the results divided into target group and categories

	Managers			Testers			Other team members		
	Average	Variance	Skewness	Average	Variance	Skewness	Average	Variance	Skewness
Planning	4,16	1,34	-1,29	3,46	1,51	-0,54	3,78	1,41	-1,06
Organizing	4,08	1,13	-1,46	3,62	1,57	-0,71	3,57	1,46	-0,70
Leading	4,10	0,86	-1,21	3,25	1,63	-0,33	3,58	1,19	-0,34
Controlling	4,01	1,31	-1,42	3,29	1,51	-0,43	3,51	1,62	-0,81

Source: own work.

We already know that there are statistically significant differences between the average responses of managers, testers and members, so we will no longer analyze whether there are significant statistical differences between the averages for groups by category. We will look at whether there are differences between the categories for each of the groups separately, because we had no reason to reject the hypothesis that the averages for the categories (without a division into groups) are equal at the given significance level. For this purpose, we will perform a one-way ANOVA tree times at the significance level $\alpha = 0,05$. The first will be performed with the following test hypotheses for managers only:

H0: $\mu_{PlanningM} = \mu_{OrganizingM} = \mu_{LeadingM} = \mu_{ControllingM}$

H1: it's not true that H0

P-value is 0,799625. It's higher than the significance level, that is, we don't have the basis to reject the zero hypothesis. On the basis of the sample, we can therefore conclude that, at the accepted significance level, for managers the differences between the averages in the populations of the categories under consideration are not statistically significant. The second analysis will be performed with the following test hypotheses for testers only:

H0: $\mu_{PlanningT} = \mu_{OrganizingT} = \mu_{LeadingT} = \mu_{ControllingT}$

H1: it's not true that H0

P-value is 0,048174. It's lower than the significance level, that is, we can reject a zero hypothesis in favor of an alternative hypothesis and conclude that for testers the differences between the averages in populations of the samples are statistically significant. The third analysis will be performed with the following test hypotheses for other team members only:

H0: $\mu_{PlanningO} = \mu_{OrganizingO} = \mu_{LeadingO} = \mu_{ControllingO}$

H1: it's not true that H0

P-value is 0,326192. It's higher than the significance level, that is, we don't have the basis to reject the zero hypothesis. On the basis of the sample, we can therefore conclude that, at the accepted significance level, for other team members the differences between the averages in the populations of the categories under consideration are not statistically significant.

Different authors distinguish management functions differently, for example by adding human resources (Knootz, O'Donnell, 1972) to them or identification (Stabryła, 2018). Others omit leading, limiting them to planning, organizing and controlling (Steinmann, Schreyogg, 2001). Some authors expand this classification even further, listing as many as nine functions: planning.

The obtained p-value for testers was very close to the significance level. An additional Tukey HSD post-hoc analysis has been carried out which showed that there are no statistically significant differences between the averages for the categories – planning and organizing (H0: $\mu_{PlanningT} = \mu_{OrganizingT}$, H1: it's not true that H0, p-value = 0,6604168), planning and leading (H0: $\mu_{PlanningT} = \mu_{LeadingT}$, H1: it's not true that H0, p-value = 0,5602368), planning and controlling (H0: $\mu_{PlanningT} = \mu_{ControllingT}$, H1: it's not true that H0, p-value = 0,7392183), organizing and leading (H0: $\mu_{OrganizingT} = \mu_{LeadingT}$, H1: it's not true that H0, p-value = 0,0567492), organizing and controlling (H0: $\mu_{OrganizingT} = \mu_{ControllingT}$, H1: it's not true that H0, p-value = 0,1653593), leading and controlling (H0: $\mu_{LeadingT} = \mu_{ControllingT}$, H1: it's not true that H0, p-value = 0,8999947).

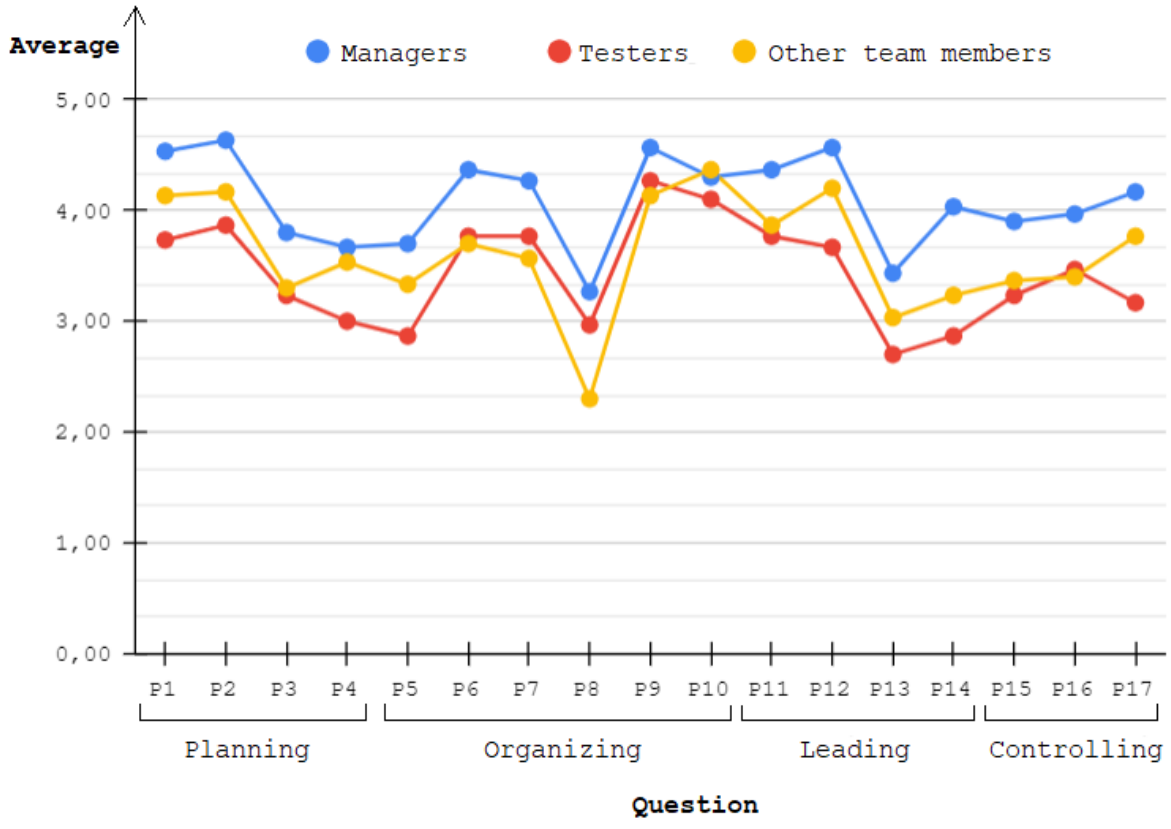

Figure 2. The chart of average results for questions by target groups. Source: own work.

We will carry out an analysis at an even deeper level, namely by target groups and by questions. The average, variance and skewness of the results are presented in a table 8 and it was a base for a chart, shown in the figure 2. Additionally, a table 10 has been prepared, in which the questions for each target group were sorted according to the average obtained in them.

The lowest rated questions were those with an average answer of less than 3, below the neutral answer. They are bolded in the table 10. There were no such questions among managers at all, there were four such questions among testers, and the other members included one. Both testers and the other members pointed to question P8, concerning the delegation of managers responsibilities and decision-making powers, if testers have the appropriate competence. Testers also rated the lowest P13 question about differentiated treatment based on their competencies and commitment, P5 question about inconsistencies between estimation of the time needed for tests and how much time actually needs to be spent on them and P14 question about motivation for testers to their work. Summing up, among the lowest rated questions were questions from the areas of organizing (P5, P8) and leading (P13, P14).

Managers rated the highest on P2 question related to defining what the tester should achieve with his actions. On the other hand, the testers rated the highest P9 question about communication between testers. The other team members rated the highest on P10 question regarding communication between testers and other team members. So it's worth to check how the communication between testers and managers was rated (P11 question) – it also doesn't look bad and in each of target groups it was in the second half of the total number of questions.

Table 8.
Average, variance and skewness of the results divided into target group and questions

		Managers			Testers			Other team members		
		Average	Variance	Skewness	Average	Variance	Skewness	Average	Variance	Skewness
Planning	P1	4,53	0,88	-2,66	3,73	1,31	-0,92	4,13	0,81	-1,80
	P2	4,63	0,65	-3,40	3,87	0,95	-0,92	4,17	0,63	-1,21
	P3	3,86	1,39	-0,73	3,23	2,05	-0,37	3,30	1,94	-0,42
	P4	3,67	1,82	-0,52	3,00	1,38	0,00	3,53	1,84	-0,74
Organizing	P5	3,70	1,32	-1,11	2,87	1,43	-0,37	3,33	1,06	-1,15
	P6	4,37	0,52	-1,29	3,77	1,15	-0,94	3,70	1,04	-0,79
	P7	4,27	0,62	-1,44	3,77	1,50	-0,73	3,57	1,29	-0,56
	P8	3,27	2,41	-0,24	2,97	1,96	-0,10	2,30	1,46	0,38
	P9	4,57	0,25	-0,28	4,27	1,03	-1,43	4,13	0,81	-0,58
	P10	4,30	0,63	-2,38	4,10	0,85	-1,05	4,37	0,65	-1,63
Planning	P11	4,37	0,45	-1,33	3,77	1,43	-0,95	3,87	0,81	-0,33
	P12	4,57	0,32	-0,88	3,67	1,26	-0,68	4,20	0,92	-1,68
	P13	3,43	1,56	-0,46	2,70	1,53	0,15	3,03	1,55	0,28
	P14	4,03	0,45	-0,04	2,87	1,57	-0,07	3,23	0,67	0,34
Controlling	P15	3,90	1,20	-1,48	3,23	1,29	-0,49	3,37	1,48	-0,90
	P16	3,97	1,41	-1,38	3,47	1,71	-0,47	3,40	1,70	-0,52
	P17	4,17	1,39	-1,70	3,17	1,59	-0,45	3,77	1,70	-1,23

Source: own work.

In the table 9 are shown questions that have statistically different averages for groups. This was calculated using the Tukey HSD test[2] at the significance level $\alpha = 0{,}05$ and only the results obtained are presented in the table. The notations used are: M – averages are significantly different for managers, T – averages are significantly different for testers, O averages are significantly different for other team members. Based on this in the table 10 have been marked with blocks questions for which we cannot exclude that their average for the whole population are equal to those previously considered to be the lowest (bold in the table 10). For example, for testers for question P13 this was done in the following way: If you selected P13 question earlier, you should also mark the P1, P2, P3, P4, P5, P7, P12, P14, P15, P16, P17 questions because there is a significant probability that differences in these averages are not coincidental. The same operation was performed for the other bold questions.

Table 9.
Statistically different averages for questions divided by target groups

	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	P13	P14	P15	P16	P17
P1	-							MO					MO				
P2		-	T	M	M			MO					MO				
P3		T	-					O		O							
P4		M		-				O	T	T							
P5		M			-			O	T	TO							
P6						-		MO					MT				
P7							-	MO					T				
P8	MO	MO	O	O	O	MO	MO	-	MTO	MTO	MO	MO			O	O	O
P9				T	T			MTO	-				MTO	T			T
P10			O	T	TO			MTO		-			TO	TO	O		
P11								MO			-		MT				
P12								MO				-	MO				
P13	MO	MO				MT	T		MTO	TO	MT	MO	-				
P14									T	TO				-			
P15								O		O					-		
P16								O								-	
P17								O	T								-

Source: own work.

Among testers only P9, P10 and P11 questions stayed unchecked – all of them relate to communication, respectively, between testers, between testers and other team members, and between testers and managers. For other team members, only two additional questions were marked, namely P13 and P14 – both of these questions were previously bolded for testers. Finally, all the questions selected, both in bold and marked rectangles, should be considered the lowest.

[2] Due to the limitation of the number of columns in the previous tool, in this case the calculation was done with: One Way ANOVA Calculator, Analysis Of Variance, Tukey HSD test, https://www.statskingdom.com/180Anova1way.html.

Table 10.
Average results for questions sorted ascending

Managers		Testers		Other team members	
Question	Average	Question	Average	Question	Average
P8	3,27	**P13**	2,70	**P8**	2,30
P13	3,43	**P5**	2,87	P13	3,03
P4	3,67	**P14**	2,87	P14	3,23
P5	3,70	**P8**	2,97	P3	3,30
P3	3,80	P4	3,00	P5	3,33
P15	3,90	P17	3,17	P15	3,37
P16	3,97	P3	3,23	P16	3,40
P14	4,03	P15	3,23	P4	3,53
P17	4,17	P16	3,47	P7	3,57
P7	4,27	P12	3,67	P6	3,70
P10	4,30	P1	3,73	P17	3,77
P6	4,37	P6	3,77	P11	3,87
P11	4,37	P7	3,77	P1	4,13
P1	4,53	P11	3,77	P9	4,13
P9	4,57	P2	3,87	P2	4,17
P12	4,57	P10	4,10	P12	4,20
P2	4,63	P9	4,27	P10	4,37

Source: own work.

5. Summary

Each of the management functions can be reflected in the management of testers. Designed and conducted survey showed that the overall condition of the testers management is not the worst. It's hard to distinguish specific management functions that would look better or worse in the study – the condition of each can be described as good but not very good, which creates place for improvement. It was clearly noted that managers were the best to judge the management of the testers, and they actually judged their own work. Worse results were reported in the assessments of the other team members, and the worst management was judged by the most involved in testing – testers. When the results were divided by the target groups and the differences in averages between the different management functions were analyzed separately for each of them, it was once again difficult to find significant differences among all the groups.

References

1. Blanchard, K., Zigarmi, P., Zigarmi, D. (2019). *Jednominutowy menedżer i przywództwo. Przywództwo sytuacyjne i lepsza komunikacja na linii przełożony-podwładny.* Warszawa: MT Biznes.
2. *Certyfikowany tester. Sylabus poziomu podstawowego ISTQB.* International Software Testing Qualifications Board, Stowarzyszenie Jakości Systemów Informatycznych, ver. 2018, V 3.1.
3. Griffin, R.W. (2004). *Podstawy zarządzania organizacjami.* Warszawa: PWN.
4. Jemielniak, D., Koźmiński, A.K. (2011). *Zarządzanie od podstaw.* Warszawa: Wolters Kluwer.
5. Kaczmarek, M., Tarka, P. (2013). *Metoda gromadzenia danych a ekwiwalencja wyników pomiaru systemu wartości w 5- i 7-stopniowych skalach ratingowych likerta.* Handel Wewnętrzny.
6. Koontz, H., O'Donnell, C. (1972). *Principles of Management: An Analysis of Managerial Functions.* New York: McGraw-Hill Companies.
7. Koźmiński, A.K., Piotrowski, W. (2007). *Zarządzanie. Teoria i praktyka.* Warszawa: PWN.
8. Paliszkiewicz, J. (2018). *Kreowanie w mediach społecznościowych wizerunku kandydata do pracy na przykładzie portalu linkedin.* Human Resource Management/Zarządzanie Zasobami Ludzkimi.
9. Roman, A. (2015). *Testowanie i jakość oprogramowania. Modele, techniki, narzędzia.* Warszawa: PWN.
10. Stabryła, A. (2018). *Podstawy organizacji i zarządzania. Podejścia i koncepcje badawcze.* Kraków: Wydawnictwo Uniwersytetu Ekonomicznego.
11. Steinmann, H., Schreyögg, G. (2001). *Zarządzanie. Podstawy kierowania przedsiębiorstwem. Koncepcje, funkcje, przykłady.* Wrocław: Oficyna Wydawnicza Politechniki Wrocławskiej.
12. Stokes, Y., Vandyk, A., Squires, J., Jacob, J.D., Gifford, W. (2019). *Using facebook and linkedin to recruit nurses for an online surve.*, Western Journal of Nursing Research.
13. Whittaker, J.A., Arbon, J., Carollo, J. (2012). *Testuj oprogramowanie jak Google. Metody automatyzacji.* Gliwice: Helion.