MASTER THESIS

Faculty of Behavioural, Management and Social Sciences

THE EFFECT OF TEAM DIVERSITY ON PERFORMANCE AND THE ROLE OF TEAM INCLUSION ON THIS RELATIONSHIP

Birgit H.M. Maas

M.Sc. Educational Science and Technology

EXAMINATION COMMITTEE

First supervisor
dr. E.M.J. Aarntzen

Second supervisor
prof. dr. M.D. Endedijk

February 11th, 2022

UNIVERSITY OF TWENTE.

Table of contents

Acknowledgement

Abstract

Diversity research in technical teams shows ambiguous results. On the one hand, more diversity can lead to better performance when more knowledge is available in such teams. On the other hand, differences between team members can lead to conflict which is detrimental for the performance. It is valuable to know how to decrease the negative and stimulate the positive interactions between team diversity and performance. This study aims to gain insight in that relationship by focussing on two types of diversity: surface-level and deep-level diversity. Surface-level diversity focusses on differences in age, gender, etc., while deep-level diversity entails differences in skills, personality, etc. Two hypotheses were formulated to structure this study and to learn how to benefit from diversity in teams. First, it is hypothesized that there is a positive relationship between deep-level diversity (i.e., professional identity) and performance. The second hypothesis states that feeling included moderates the surface-level diversity (i.e., gender and nationality) – performance relationship. This means that when team members feel included, surface-level diversity will not have a negative effect on the performance of the team. Data collection took place at a university of applied science in a project week with 71 teams. With two online questionnaires (one on the first day and one on the fourth day of the project week), data were collected about team diversity (i.e., gender, nationality, and professional identity), performance, and inclusion. The first hypothesis was answered with a regression analysis that showed an insignificant relationship between deep-level diversity and performance. For the second hypothesis, a significant main effect between gender diversity and performance was found. Moreover, a moderation analysis showed that inclusion moderated the gender diversity – performance link. When team members felt included, gender diversity did not have an effect on performance, when team members did not feel included, more gender diversity led to higher performance. No significant relationships were found with nationality. Lastly, a positive relationship was found between inclusion and subjective performance. In summary, inclusion moderated the gender diversity – subjective performance link. Also, inclusion and subjective performance are positively related. Therefore, it is concluded that gender diversity and inclusion are key elements for good team performance. For future research, this study suggests a larger sample size, more variety in surface-level diversity, and a longitudinal study to collect data over a longer period to get a better and more complete insight in the diversity – performance relationship.

Introduction

Technical teams are often diverse in educational background, experiences, and disciplines. Diversity is defined by Knippenberg, Ginkel, and Homan (2013) as differences between team members in gender, age, cultural background, tenure, educational background, and functional background. Literature shows that diverse (heterogeneous) teams have more knowledge available, while homogeneous teams often have less collaboration problems (e.g., communication issues), which both can lead to better performance (Van Knippenberg et al., 2004). These two contradicting perspectives can explain the inconsistent effects of diversity on performance that have been found in the literature. Since diverse teams have the potential to collaborate well under the right circumstances, the question remains under which circumstances diverse teams in a technical sector can collaborate fluently.

In the current study, it is expected that diverse teams will have better performance than homogeneous teams, but only when team members feel included in a team. Diverse teams have the possibility to perform better since such teams may learn more from the combination of perspectives, networks, and knowledge available. However, because of the social categorization processes (i.e., conflict between team members because of ingroup and outgroup) team members may identify less with their team and therefore diverse teams may perform worse. Feeling included in a team could lead to higher team identification and less conflict between team members and thus increase team performance. Additionally, Harrison et al. (2002) found that feeling included was a strong predictor of team performance.

This study builds on the work of Van Veelen and Ufkes (2019). They found that team diversity facilitated performance because of increased team learning and team efficacy, but only for teams with a strong team identity. For teams with a weak team identity, team efficacy only had a positive effect on subjective and not objective performance. Also, the potential for learning and performance of a diverse team is not present in teams with a weak group identity. Additionally, they found that objective diversity seemed a better predictor of performance especially for strongly identifying teams. Objective diversity means the actual diversity in teams rather than the perceived (subjective) diversity. The research of Van Veelen and Ufkes focusses on surface-level diversity (i.e., observable differences, for example, age or gender). This study adds deep-level diversity (i.e., less observable differences, for example, values or expertise) since, based on literature, it is expected that team performance can benefit from this type of diversity. This study is unique since this combination of surface-level diversity, deep-level diversity, performance, and inclusion has not been studied so far. Since deep-level

diversity is not visible, it is assumed that these differences will not cause social categorization. So, it was decided not to look at the effect of inclusion on the deep-level diversity – performance link. So, this research is innovative and relevant to increase knowledge and awareness about these aspects and their relation to each other. Finally, similar to the study of Van Veelen & Ufkes (2019), next to measures of perceived performance, also objective performance was measured to get a comprehensive insight in the effect of diversity on performance. Also, this will avoid bias as Horwitz and Horwitz (2007) found that people who are rating their own work instead their work being rated, tend to boost their own ratings.

To summarize, this study aims to explain to what extend deep-level diversity affects performance. Additionally, it aims to examine the effect of the feeling of inclusion on the surface-level diversity – performance relationship. The study took place among students who worked in multidisciplinary teams in a technical context.

Theoretical Framework

How Diversity Affects Performance

As mentioned by Van Knippenberg et al. (2004), the effects of diversity on performance are inconsistent. On the one hand, interprofessional or diverse teams have more informational resources and performance can be improved by the different perspectives, networks, and knowledge team members bring to the table. On the other hand, there might be collaboration problems because of those differences between team members. In the theoretical model of Van Knippenberg et al. (2004) a distinction is made between these perspectives, they are called the information-elaboration perspective and the social categorization perspective. The informational elaboration focusses on the task-related aspects of the teamwork and it entails that diverse team members might have more task-related knowledge. This can lead to creative and innovative ideas, which can increase performance, but can also possibly increase task conflict (Van Knippenberg et al., 2004). It is important to remember that teams with a variety of knowledge available will have to share their information before it leads to better performance. However, Van Knippenberg et al. (2004) continues that information sharing does not happen automatically and that discussing task-relevant information in a team is essential to benefit from diversity and perform well. So, not the availability of knowledge, but the use of it enhances team performance.

Social categorization on the other hand, focusses on the relational aspects of the teamwork and it entails that differences between team members can result in ingroups and outgroups. The feeling that a member belongs in a group, increases group cohesion, which means less relational difficulties, less members who want to leave the team, and higher performance (Nishii, 2013). However, there can be division in a team as well, for example when team members respond better to members in their subgroup than outside that group, this is called intergroup bias (Van Knippenberg et al., 2004). Threatening or challenging the subgroup identity might drive intergroup bias, which is disturbing for collaboration, especially in heterogeneous groups (Van Knippenberg et al., 2004). For example, in a gender diverse team that must collaborate on a mathematics problem, but the input of women is not taken seriously, this hinders optimal team performance. So, it is essential that team members share their expertise and respect each other's expertise to enhance team performance.

Van Knippenberg et al. (2004) combined the information-elaboration and social categorization perspectives in the categorization-elaboration model (CEM). The CEM was used as a start to formulate a theory that underlines the importance of team learning and team efficacy

as mediators and team identification as moderator to determine how and when diversity facilitates performance. Since the information elaboration and social categorization perspectives interact during teamwork, it is important to stimulate the information-elaboration side and diminish the negative side of the social-categorization perspective. Therefore, stimulating discussion, strengthening the subgroup identity, and removing challenges or threats to the subgroup identity might be options to improve performance in heterogeneous teams. Also, reducing threats and challenges to the subgroup identity by increasing inclusion could diminish intergroup bias.

This thesis is based upon the CEM model and thus assumes that heterogeneous groups have the potential to perform well under certain circumstances. Additionally, it is proposed that the relationship between diversity and performance depends on the different types of diversity. It is expected that with visible/ surface-level diversity (e.g., gender, nationality) the social categorization processes may be stronger than with invisible/ deep-level diversity (e.g., values, expertise). Whereas, with deep-level diversity the information elaboration process may be stronger. Therefore, it is assumed that the link between diversity and performance is much stronger for deep-level diversity since this is related to the task related aspects of teamwork.

Differences between Deep-level Diversity and Surface-level Diversity and their Relationships with Performance

Deep-level diversity involves the less observable aspects (e.g., personality, attitudes, values, expertise, function, ability, racial/ethnic /identity, sexual identity, and other social identities), while surface-level diversity involves observable differences (e.g., gender, age, nationality, ethnicity, and religion) (Jackson & Joshi, 2011; Van Knippenberg et al., 2004; Van Veelen & Ufkes, 2019). These types of diversity seem to be related differently to performance. Horwitz and Horwitz (2007) found that deep-level diversity was significantly positively related to the quality of team performance. Deep-level diversity in this study was about individual attributes that are relevant for accomplishing tasks (e.g., functional expertise, education, and organizational tenure). Additionally, it has been found that the positive effects of deep-level diversity are increased over time (Harrison et al., 1998). On the other hand, no significant relationship was found between surface-level diversity and team performance (Horwitz & Horwitz, 2007). Though, the negative relationship between surface-level diversity and team integration/ cohesion diminished over time (Harrison et al., 1998).

So, surface-level diversity potentially leads to worse performance. It is expected that the observable differences between team members can distract from the teamwork and eventually lead to subgroups within a team. These aspects can be linked with the social categorization perspective. The differences between team members can lead to ingroups and outgroups within a team, which impedes team performance. In this study, surface-level diversity in this study is measured by gender and nationality. In the Netherlands, in science, technology, engineering, and mathematics (STEM) studies and work field women are underrepresented (Booy et al., 2012). This means that gender will be a salient characteristic in the teams. The meta analysis of Tshetshema and Chan (2020) noted that 10 out of 14 articles found a positive relationship between gender diversity and performance. So, it seems to be beneficial to have gender diverse teams. Additionally, this study takes place in the Netherlands, so most of the respondents will be Dutch. This means that nationality will be a salient characteristic in the teams. Tsui et al. (1992) found that high nationality diversity in a team leads to less commitment and attachment to an organization and to more absence. Moreover, when a team includes non-Dutch speaking members, they will have to communicate in English which might lead to communication problems. Larkey (1996) proposes that majority members stick with their communication patterns while minority members will match the communication patterns of the majority members. So, it is interesting to investigate what the effect is of nationality diversity on performance. Moreover, gender and nationality differences in a team have more effect on attachment to an organization (Tsui et al., 1992) and have stronger relationships (Nishii, 2013) than the other variables (e.g., age, tenure, education). Therefore, in this study, surface-level diversity is measured by gender and nationality.

Deep-level diversity on the other hand leads to good performance. It is expected that since there is knowledge and expertise to share, this facilitates team performance. This quality can be associated with the information elaboration perspective that suggests that sharing knowledge and discussing it results in good team performance. In this study, deep-level diversity is measured by the professional identity of team members. It is expected that high PI diversity leads to better performance since such teams have a wider variety of knowledge and skills than teams with less PI diversity. The aspects to measure professional identity are interest, values, competencies, and characteristics. These are also aspects of deep-level diversity, therefore, professional identity is suitable to measure deep-level diversity.

Deep-level diversity and surface-level diversity have quite a different relationship with team performance. It is valuable to see if there is a situation that diminishes the negative effect of surface-level diversity in the surface-level diversity – performance link. To ensure that team

members do not focus mainly on their differences, so to prevent social categorization, it could help if they feel included in their team. When team members can focus on the project because they are not distracted by differences and they dare to speak their minds because they feel included, this should lead to better performance. Also, interest, values, competencies, and characteristics are less noticeable at first sight and therefore there is a smaller chance that the social categorization processes are activated. So, it is expected that deep-level diversity results in higher performance, regardless of the perceived team inclusion.

Based on this literature that was studied, the following research question and hypothesis were formulated:

RQ1: What is the role of deep-level diversity (i.e., professional identity) on performance?

Hypothesis 1: Deep-level diversity is associated with better performance.

Moderating Effect of Inclusion

Deep-level and surface-level diversity are expected to have different relationships with team performance. To minimize the negative effect of surface-level diversity (caused by social categorization) on performance, inclusion was suggested as moderator. Inclusion is defined by Jansen et al. (2014) as "the degree to which an individual perceives that the group provides him or her with a sense of belonging and authenticity" (p. 373). Based on the CEM, there are two reasons why inclusion leads to better team performance. First, when team members feel included there may be less focus on their differences, so there is less distinction between ingroup and outgroup (i.e., intergroup bias), which is beneficial for collaboration. For example, when there is a Dutch ingroup and a non-Dutch outgroup and people feel that ideas of non-Dutch people in the team are not as valuable as the ideas from others, it is possible that the team is not performing optimally. Second, it is hypothesized that when team members feel included in a team, it might be easier to participate and contribute to the task related discussion (Bodla et al., 2018) which is beneficial for the performance of a team. In line with these expectations, Bodla et al. (2018) found that feeling included diminishes the negative effects of surface-level diversity and that the feeling of inclusion moderates the perceived diversity – collaboration relationship. Additionally, Harrison et al. (2002) state that inclusion was a strong predictor for performance. More specifically for gender, Nishii (2013) found that inclusive climate moderates the relationship between gender diversity and task conflict. So that gender diverse teams in an inclusive climate have less conflict, which leads to better performance (De Dreu & Weingart, 2003).

Based on these findings in the literature that was studied, the following research question and hypothesis were formulated:

RQ2: What is the moderating effect of feeling included on the relationship between surface-level diversity (i.e., gender and nationality) and performance?

Hypothesis 2: The effect of surface-level diversity is associated with lower levels of performance but moderated by feeling included in a team; the higher team inclusion, the weaker the negative effect of surface-level diversity on performance.

The Current Study

In this study, the effect of surface-level and deep-level team diversity (i.e., gender, nationality, and professional identity) on performance (subjective and objective) was investigated. Additionally, the effect of feeling included within a team was examined to see if it moderated the surface-level diversity – performance relationship (see Figure 1). The data were collected during a project week at a University of Applied Sciences in the Netherlands. The students in this project week had different study backgrounds (e.g., electrical engineering, industrial product design, mechanical engineering) and they were assigned to various teams. Gender and nationality were notable aspects of participants since only 14% was woman and 17.7% had the German nationality (while the other 82.3 % was Dutch). To make team members aware of the deep-level diversity in their team, all team members participated in an exercise measuring their professional identity which then had to be discussed within the team.

Figure 1

Research model
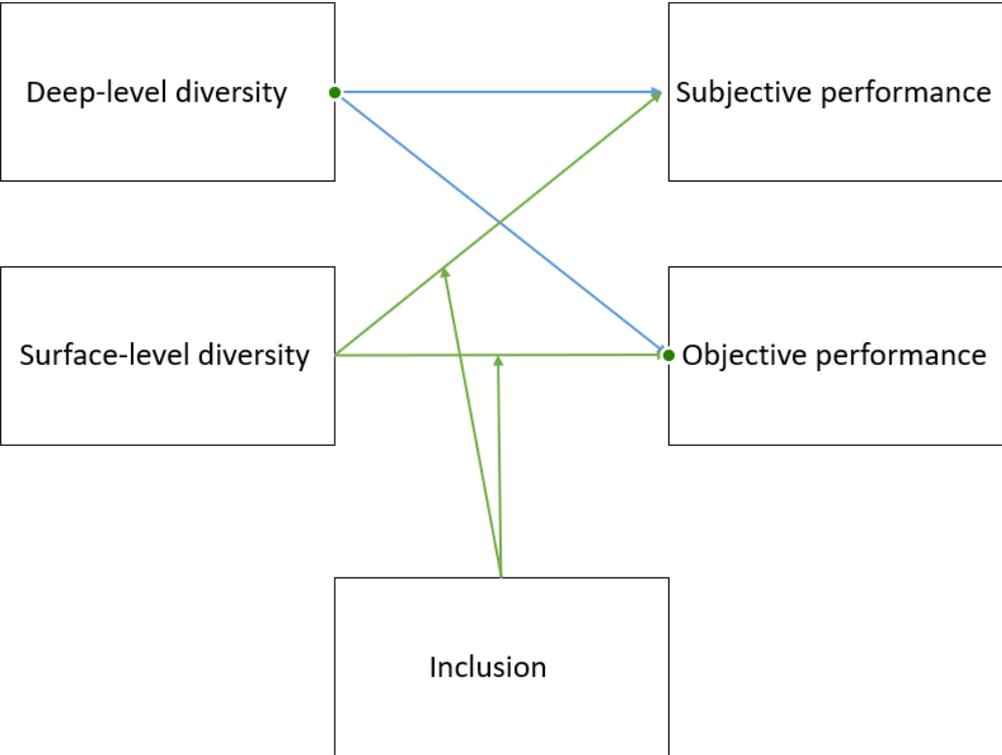
Method

Research Design

This was a correlational study to examine the relationship between diverse teams (i.e., diverse in *gender*, *nationality*, and *professional identity*) and *performance* and the moderating effect of *inclusion* on this relationship.

Respondents and Procedure

Respondents

The participants in this study were all students who followed the international project week organized by a University of Applied Sciences in The Netherlands. In this week, the participants had to work together for four days on a problem or task from a company. They received their task on a Monday morning and had to present their results on Friday. During the week, the teams would come together every day to work on and develop their solution(s), idea(s), and/ or product(s). The teams had 5-8 members who mostly did not know each other beforehand and they had different educational backgrounds and nationalities.

The participants were students from the life science, engineering, and design (LED) studies in their first, second, third, or fourth year. Additionally, there were international exchange students from different countries and last year secondary school students (i.e., Havo-5; students who wanted to get to know the University of Applied Sciences to help with the decision if they wanted to follow a study program there). The team leaders of the teams were fourth- or fifth-year LED students who participated in the international project week at least three times before.

This research was part of a larger study[1]. For this research only part of the data were used (i.e., group in which the professional identity of all team members was made salient), therefore the procedure of the control group will not be reported.

Procedure

First Questionnaire. The first day of the international project week started with an introduction for the participants about the project week in general and a presentation about the

---

[1] In the original study, the participants were divided in an experimental and a control group. Both groups filled out three questionnaires and the experimental group did an exercise as well. For this study only the data of the experimental group were analysed.

project they would be working on during that week. After the presentations, they went to their working space. Here this study started. There was a general explanation about the questionnaire, that measured demographics, and the exercise to prime team members of each other's professional identity diversity which had to be done before the teams could start working on their project.

Professional Identity Exercise. This professional identity exercise was included to be sure that this form of *deep-level diversity* was salient for all participants even though they did not know each other well yet. A link to the first questionnaire (that was programmed in Qualtrics; see Appendix A) was shared and filled in by the participants individually. Also, paper instructions (Appendix C; Figure 8 and Figure 9) and worksheets (Appendix C; Figure 10, Figure 11, and Figure 12) were distributed among the teams for the exercise. At the end of the survey, team members were directed to the beginning of the exercise. For this exercise, participants had to answer statements (programmed in Excel; Appendix E) about their interests, competencies, values, and personality which resulted in scores for five professional profiles (i.e., All-rounder, Analyst, Individualist, Innovator, Team player) (for more details, see measures). Subsequently, team members gained insight in the five different profiles and their characteristics with the descriptions of Endedijk et al. (2019) (Appendix C; Figure 11). Then, all team members picked a final profile which were then combined on the team board (Appendix C; Figure 12). In the last step of the exercise, all team members explained the choice of their profile, so everyone gained insight in the professional identity of all team members. To stimulate participation and thank the participants, all teams received a cake when they were finished with the first survey and the exercise.

Second Questionnaire. On the fourth day of the project week, all teams received a link to the second questionnaire (also programmed in Qualtrics; Appendix B), in which *subjective performance* and *inclusion* were measured. The teams also received an explanation on paper that this is still part of the research. At the end of the questionnaire there was a code that could be filled out on the paper and brought to a central point where they could get coins for all team members to pick up a free snack in the canteen. This was done to stimulate participation and to thank them for their cooperation.

Complete randomization was not possible in this study since the teams were already formed and the locations where they had to work were fixed, therefore, a convenience sample was used.

There were 521 responses to the first questionnaire and 490 to the second questionnaire. The dataset that was used for analyses contained 379 responses. Responses that were deleted were from participants who indicated that they did not want their data to be used for analysis, double responses, empty responses, test responses, and incomplete responses. So, in total, 379 students (14.0 % female) in 71 teams participated in this study. Their age varied from 16 to 30 ($M =$ 20.56; $SD =$ 2.24). For further descriptive characteristics of the participants (i.e., nationality, educational institution, study program), see Table 1.

Table 1

Descriptive information respondents

	n	%
Gender		
Male	326	86
Female	53	14
Nationality		
Dutch	312	82.3
German	67	17.7
Educational institution		
University of Applied Sciences	334	88.1
Dutch secondary education	7	1.8
Other	38	10
Study program		
Applied physics	31	8.2
Chemical technology	7	1.8
Chemistry	9	2.4
Electrical engineering	29	7.7
Industrial product design	37	9.8
Mechanical engineering	110	29.0
Mechatronics	78	20.6
Technical business	7	1.8
Technical computer science	40	10.6
Other	24	6.3
Study year		

1	150	39.6
2	118	31.1
3	70	18.5
4	29	7.7
5 or higher	5	1.3

Ethical Issues

Before the data collection started, an ethics request was sent to the ethics committee of the University of Twente and got approved. The informed consent was taken care of in the first questionnaire. The University of Applied Sciences helped with getting as many responses as possible by making the questionnaires a mandatory part of the project week. However, the participants could indicate in the first and second questionnaire that they did not want their responses analysed. In both cases, the responses were left out of the data analysis. The data was anonymized, and personal data was deleted from the file before it was used for analysis. After a few weeks, when all mail addresses were collected, an email was sent to all participants to debrief them about the study they took part in.

Measures

The data were collected with two questionnaires at two moments in time in the international project week. The first questionnaire (i.e., conducted on the first day of the project week) measured demographics and diversity (i.e., *gender*, *nationality*, and *professional identity*) and the second questionnaire (i.e., conducted on the fourth day of the project week) measured the *subjective performance* and the *feeling of inclusion*. The *objective performance* was received later via mail. Table 2 gives an overview of when the variables relevant for this study were measured[2]. Each measure is explained in detail after Table 2.

---

[2] For an overview of all measurements, please contact the examination committee.

Table 2

Measurement of variables in multiple surveys

	Diversity	Performance	Inclusion
Survey 1 (day one in project week)	X		
Survey 2 (day four in project week)		X	X

Gender Diversity in Teams

To calculate the *gender diversity* in a team, Blau's heterogeneity index (Blau, 1977) was used. The formula to calculate the heterogeneity index, is $(k/k-1)*(1-\sum_{i=1}^{R} p_i^2)$, with k being the number of categories (two in the case of gender, i.e., men/women). The heterogeneity index gets a value between 0 and 1 with 0 being no diversity, everyone belongs in the same category and 1 being maximum diversity when a team has an equal number of men and women. For example, if a team had four men and one woman, the calculation would be $(2/1) * 1 - (0.8^2 + 0.2^2) = 1-(0.64+0.04) = 2*(1-0.68) = 0.64$. So, the heterogeneity for this team would be 0.64. In case there are more women than men in a team the diversity index is calculated in the same way. So, a team with four men and one woman has the same index as a team with four women and one man. Table 3 shows a summary of the gender heterogeneity index. In this study, Blau's heterogeneity index varied between .00 and .96 (*M* = .38; *SD* = .38) and there were two teams that counted more women than men. Table 4 shows the percentages of women in the teams and the first row shows all teams with only men.

Table 3

Gender diversity in teams (Blau's index)

Gender diversity	Number of teams (%)
Very low diversity (< .20)	34 (47.9)
Low diversity (.20-.39)	0 (0)
Moderate diversity (.40-.59)	9 (12.7)
High diversity (.60-.79)	14 (19.7)
Very high diversity (≥.80)	14 (19.7)

Table 4

Percentages women in teams

% women in teams	Number of teams (%)
0	34 (47.9)
1-20	18 (25.4)
21-40	17 (23.9)
41-60	0 (0)
61-80	2 (2.8)
81-100	0 (0)

Nationality Diversity in Teams

To calculate the *nationality diversity* in a team, also Blau's heterogeneity index (Blau, 1977) was used (see '*Gender diversity in teams*' for explanation). This calculation also does not distinguish between more Dutch or more non-Dutch people in a team, same as for the gender diversity calculations. This study, there were only three teams that counted more non-Dutch than Dutch people. Table 5 shows a summary of the nationality heterogeneity index. The Blau's heterogeneity index varied between .00 and 1.00 ($M = .41$; $SD = .44$) and there were three teams that had more non-Dutch than Dutch members. Table 6 shows the percentages of non-Dutch in the teams in this study and the first row indicates all teams with only Dutch members.

Table 5

Nationality diversity in teams (Blau's index)

Nationality diversity	Number of teams (%)
Very low diversity (< .20)	37 (52.1)
Low diversity (.20-.39)	0 (0)
Moderate diversity (.40-.59)	2 (2.8)
High diversity (.60-.79)	9 (12.7)
Very high diversity (≥.80)	23 (32.4)

Table 6

Percentages non-Dutch members in teams

% non-Dutch members in teams	Number of teams (%)
0	37 (52.1)
1-20	7 (9.9)
21-40	16 (22.5)
41-60	9 (12.7)
61-80	2 (2.8)
81-100	0 (0)

Professional Identity Diversity

The *professional identity* is based on the outcome of a shortened version of the Career Compass. This instrument was developed by Möwes (2016) and produces a professional profile (e.g., All-rounder, Analyst, Individualist, Innovator, or Team player) when all 100 items of the questionnaire are answered. Those items are divided in four categories: interests, competencies, values, and personality. Since the original version of the Career Compass is quite long and in this research it was only a part of the exercise, the choice was made to reduce the number of items to 26 (for development details and validation, see Appendix D). When this shortened version is filled in, a professional profile is produced in the same way as the original Career Compass. It depends how the items are answered which profile is the best fit. For example, if someone got a result that Team player was the profile with the best fit, they probably have characteristics like valuing relationships (e.g., family, friends, work – life balance) and being competent in managing people and team collaboration.

The *professional identity diversity* within a team was calculated with Blau's heterogeneity index as well. The formula is still $(k/k-1)*(1-\sum_{i=1}^{R} p_i^2)$, however there are five profiles (i.e., All-rounder, Analyst, Individualist, Innovator, and Team player), therefore k = 5. The heterogeneity index gets a value between 0 and 1 with 0 being no diversity, everyone had the same profile and 1 being maximum diversity when a team has an equal number of members for every profile. For example, if a team of eight members has three members with an analyst profile, two members with an allrounder profile, two members with an individualist profile, and one member with an innovator, the calculation would be $(5/4)*(1-(0.25^2 + 0.25^2 + 0.25^2 + 0.125^2 + 0.125^2)) = 1.25*(1-(0.0625 + 0.015625)) = 1.25*(1-(0.1875 + 0.03125)) = 1.25*(1-$

0.21875) = 0.97. Table 7 shows a summary of the *deep-level diversity* in the teams in this study. Blau's heterogeneity index varied between .31 and .97 ($M = .70$; $SD = .16$).

Table 7

Professional identity diversity in teams

Professional identity diversity	Number of teams (%)
Very low diversity (< .20)	0 (0)
Low diversity (.20-.39)	3 (4.2)
Moderate diversity (.40-.59)	15 (21.1)
High diversity (.60-.79)	23 (32.4)
Very high diversity (≥.80)	30 (42.3)

Subjective Performance

To measure *subjective performance*, Hackman's (1987) team effectiveness scale was adapted. The scale was originally measured by nine items, for example, 'In your opinion how well does your project team perform? Please grade your project team's performance on... - ... the overall performance level'. All items were answered on a 5-point Likert-type scale (1 = terrible, 5 = excellent). Factor analysis was performed and showed that the items could be explained by two factors. Four items measured the result (e.g., 'In your opinion how well does your project team perform? Please grade your project team's performance on... - ... accomplishing project goals') while the other five items measured the collaborative process (e.g., 'In your opinion how well does your project team perform? Please grade your project team's performance on... - ... the efficiency in collaborating'). Therefore, only the four items measuring result were chosen and used to create a subjective performance scale (i.e., factor loadings ranged from .613 to .785). Cronbach's alpha was .852.

Objective Performance

The *objective performance* was indicated with a grade by company supervisors and teachers. The grade could range between 1 (bad) and 10 (excellent). The teams performed well with an average of 7.62 as their final grade.

Inclusion

Inclusion means that team members feel accepted and equal, are valued, and establish personal connections within a team (Hui & Farnham, 2016; Jansen et al., 2014; Shore et al., 2011). The items to measure *inclusion* were adapted from Jansen et al. (2014). *Inclusion* was originally measured with six items, for example, 'What are your reflections about how the team members are included in the project assignment this week? In this team... - I think that all team members feel included.'. All items were answered on a 5-point Likert scale (1 = strongly disagree, 5 = strongly agree). Factor analysis was performed and showed that not all six items belonged to the factor inclusion (i.e., factor loadings of two items were -.117 and -.204). When those two items were removed, the remaining four belonged to one factor (i.e., factor loadings ranged from .654 to .839). Additionally, Cronbach's alpha was calculated for the reliability of the factor (i.e., .802).

Data Analysis

ICC was calculated to check whether the perception of individual team members showed enough overlap for the data to be aggregated for the analysis. The ICCs were .86 for *subjective performance* and .78 for *inclusion*. This means that 86% and 78% of the variance in the individual ratings can be explained by differences between teams and 14% and 22% respectively by differences between individuals. Such high ICC's justify aggregation to the team level, since the interrater agreement was very high (Koo & Li, 2016).

Data Aggregation

The individual responses about *feeling of inclusion* and *subjective performance* were aggregated to the team level. This resulted in one score per team for *feeling of inclusion* and one score per team for *subjective performance*. Please note that *team diversity* and *objective performance* were already measured on the team level (i.e., respectively Blau's heterogeneity index of gender, nationality, professional identity, and the grade a team received).

Analytic Strategy

First, the descriptive statistics and correlations between all study variables and important demographics were calculated. Subsequently, to test hypothesis 1 'deep-level diversity is associated with better performance', two separate regressions were used. The first regression

was *deep-level diversity* as independent variable and *subjective performance* as dependent variable. The second regression was *deep-level diversity* as independent variable and *objective performance* as dependent variable. The second hypothesis is 'the effect of surface-level diversity is associated with lower levels of performance but moderated by feeling included in a team; the higher team inclusion, the weaker the negative effect of surface-level diversity on performance'. To test this hypothesis, the main effects of *surface-level diversity* (i.e., *gender* and *nationality diversity*) on *performance* (i.e., *subjective* and *objective performance*) were examined. Then, the moderating role of *team inclusion* was examined with moderation analyses applying model 1 of the PROCESS macro of Hayes, using 5000 bootstrapped samples (Hayes, 2013). This moderation model was run four times with one independent and dependent variable at a time. The independent variables were *gender diversity* and *nationality diversity*, the dependent variables were *subjective* and *objective performance* and, in all models, *inclusion* was used as moderator. Additionally, marginally significant correlations were examined with simple slope analysis provided by Dawson (n.d.).

Results

Descriptive Statistics and Correlations

The model variables and demographics (i.e., age, amount of known team members) with corresponding means (*M*), standard deviations (*SD*), and correlations are shown in Table 8. This table shows a significant correlation between age and *nationality diversity* ($r = .55$, $p < .01$), which means that the higher the average age in a team, the higher *nationality diversity* is in that team. However, age was not correlated with any of the dependent variables, so it is unlikely that age is an alternative explanation when significant correlations are found during the analyses, therefore age was not included as covariate in the analyses.

Overall, team members rated their *performance* and their *feeling of inclusion* quite high (4.02 and 4.17 respectively on a scale of 1-5). Moreover, teams are more diverse on *deep-level* (i.e., *professional identity)* rather than *surface-level* (i.e., *gender* and *nationality*). Lastly, teams performed quite well with an average of 7.62 (on a scale of 1-10) as their final grade.

Subjective and *objective performance* were measured to gain more insight in overall *performance*. It was expected that both variables would be related, however Table 8 shows that *subjective performance* had higher correlations with other variables than *objective performance* and they were not significantly correlated with each other. Also, only *subjective performance*, but not *objective performance*, shows a significant moderate positive correlation with *inclusion*, $r = .66$, $p < .01$, which means that the higher the average *inclusion in a team*, the higher the team members rate their *performance*.

Table 8

Descriptive statistics and correlations of model variables

	M	SD	1	2	3	4	5	6	7	8
1. Age	20.54	1.07	1							
2. Known team members	1.70	0.82	-.129	1						
3. Subjective performance	4.02	0.37	.100	.093	1					
4. Objective performance	7.62	0.93	.080	.181	.131	1				
5. Gender diversity	0.38	0.38	-.215	.148	.140	.110	1			
6. Nationality diversity	0.41	0.44	.545*	-.020	.208	.014	.081	1		
7. Professional identity diversity	0.70	0.16	.010	.065	-.165	-.013	-.030	-.021	1	
8. Inclusion	4.17	0.36	-.022	.044	.656*	.181	.157	-.043	-.099	1

* Correlation is significant at the $p < .01$ level, two-tailed.

Hypotheses Testing

Hypothesis 1: Deep-level Diversity is Associated with Higher Subjective and Objective Performance.

Hypothesis 1 stated that *deep-level diversity* is associated with better *subjective* and *objective performance*. To test this hypothesis, two separate linear regression analyses were done, see Figure 2 for an overview of the results (unstandardized β coefficient). First a regression with *deep-level diversity* on *subjective performance*. In contrast to the expectation, no significant relation was found ($t = -1.39$, $p = .17$). Second, a regression with *deep-level diversity* and *objective performance* was done. Again, no significant relation was found ($t = -.10$, $p = .92$). Therefore, hypothesis 1 could not be accepted.

Hypothesis 2: Is Surface-level Diversity Associated with Worse Performance Depending on Team Inclusion?

Hypothesis 2 stated that the effect of *surface-level diversity* on *performance* depends on *team inclusion*; the higher *team inclusion*, the less negative effect of *surface-level diversity* on *performance*.

Gender Diversity. First, the main effect of *gender diversity* on *subjective performance* was analysed, additionally, the moderating effect of *feeling included* on this relationship was tested (see Figure 3 for an overview of the results). In contrast with the expectations, a marginally significant relationship was found between *gender diversity* and *subjective performance* ($\beta = 2.29$, $se = 1.16$, $t = 1.98$, $p = .05$, $LLCI = -.01$, $ULCI = 4.60$) with an R^2 of .46. This implies that when a team is more diverse in gender, they rate their performance higher. Interestingly, a significant main effect of *inclusion* on *subjective performance* was found ($\beta = .82$, $se = .12$, $t = 6.84$, $p = .00$, $LLCI = .58$, $ULCI = 1.06$), indicating that the more team members *feel included* the better they think their team performs. Then, moderation analysis was done to see whether *inclusion* moderated the link between *gender diversity* and *subjective performance*. A marginally significant moderation effect was found ($\beta = -.54$, $se = .28$, $t = -1.95$, $p = .06$, $LLCI = -1.09$, $ULCI = .01$). Since the *p*-value of the moderation relationship was marginally significant ($p = .06$), this relationship got explored by plotting the regression lines to explore the meaning of this effect with Dawson (n.d.). Figure 4 shows that when experienced *team inclusion* is low, higher *gender diversity* is associated with higher *subjective team performance*, when experienced *team inclusion* is high, *gender diversity* does not seem to predict *subjective*

team performance. This is different than what was expected since it was hypothesized that when *team inclusion* was high, higher *gender diversity* would be associated with higher *performance*.

Additionally, to see if these significant relationships are caused by more women in a team or the diversity in a team, an extra moderation analysis was done. Instead of using Blau's heterogeneity index, the number of women in a team was the independent variable. So, the main effect of *the number of women in a team* on *subjective performance* was analysed, additionally, the moderating effect of *feeling included* on this relationship was tested. A marginally significant relationship was found between *number of women in a team* and *subjective performance* ($\beta = 6.11$, *se* = 3.23, *t* = 1.90, *p* = .06, *LLCI* = -.32, *ULCI* = 12.55). This means that when a team has more women, a team rates their *performance* higher. Also, a main significant effect of *inclusion* on *subjective performance* was found ($\beta = .81$, *se* = .12, *t* = 6.96, *p* = .00, *LLCI* = .58, *ULCI* = 1.04), which implies that the more team members *feel included* the higher they rate their *performance*. Then, the moderation analysis was done to see whether *inclusion* moderated the link between *number of women in a team* and *subjective performance*. Again, a marginally significant relationship was found ($\beta = -1.45$, *se* = .76, *t* = -1.90, *p* = .06, *LLCI* = -2. 89, *ULCI* = .07) with an R^2 of .46. Since the *p*-value of the moderation relationship was marginally significant (*p* = .06), this relationship got explored by plotting the regression lines to explore the meaning of this effect (Dawson, n.d.). Figure 5 shows that when experienced *team inclusion* is low, higher *number of women in a team* is associated with higher *subjective team performance*, when experienced *team inclusion* is high, the *number of women in a team* does not seem to predict *subjective team performance*.

Second, the main effect of *gender diversity* on *objective performance* was analysed and the moderating effect of *feeling included* on this relationship was tested. The relationship between *gender diversity* and *objective performance* was not significant (*p* = .43). Also, no significant main effect of *inclusion* on *objective performance* was found (*p* = .13). And no significant moderation effect was found (*p* = .46) for this relationship.

Nationality Diversity. To start, the main effect of *nationality diversity* on *subjective performance* was tested, further, the moderating effect of *team inclusion* on this relationship was analysed (see Figure 6 for an overview of the results). No significant relation was found between *nationality diversity* and *subjective performance* (*p* = .60). But a significant main effect of *inclusion* on *subjective performance* was found ($\beta = .63$, *se* = .12, *t* = 5.23, *p* = .00, *LLCI* = .39, *ULCI* = .87). This means, that the more team members *feel included*, the higher they

indicate their *performance*. Lastly, no significant moderation effect was found with *subjective performance* (*p* = .46).

Then, the main effect of *nationality diversity* on *objective performance* was tested and the moderating effect of *team inclusion* on this relationship was analysed. The relationship between *nationality diversity* and *objective performance* was not significant (*p* = .96). Also, no significant relation was found between *inclusion* and *objective performance* (*p* = .28). And no significant moderation effect was found (*p* = .96).

Hypothesis 2 cannot be confirmed since the significant moderation effect of *gender diversity* could not be reproduced with *nationality diversity*. So, there is no unambiguous result regarding *surface-level diversity* in general.

Figure 2

Regression coefficients for the deep-level diversity – performance relationship
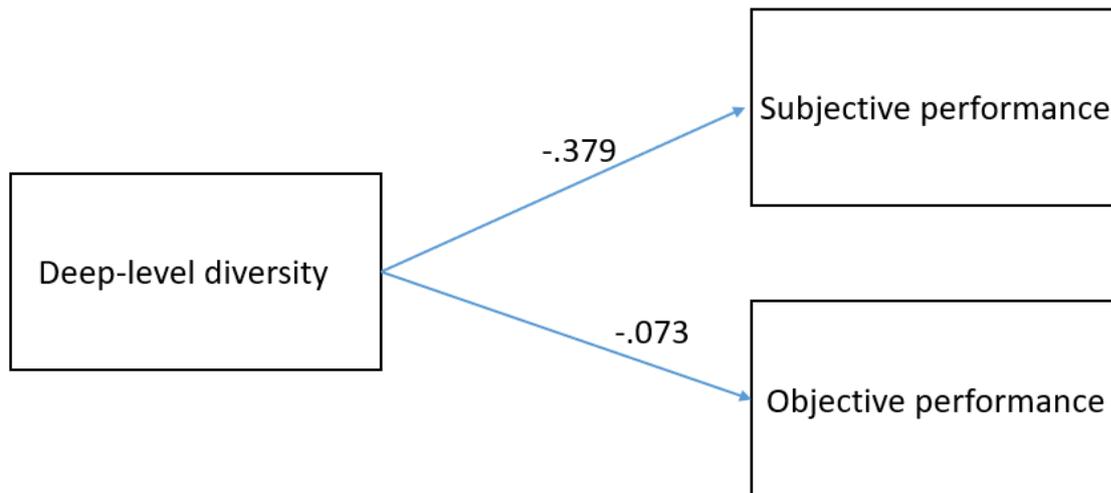
Note. No significant effects were found.

Figure 3

Moderation coefficients for gender diversity – performance relationship with inclusion as moderator
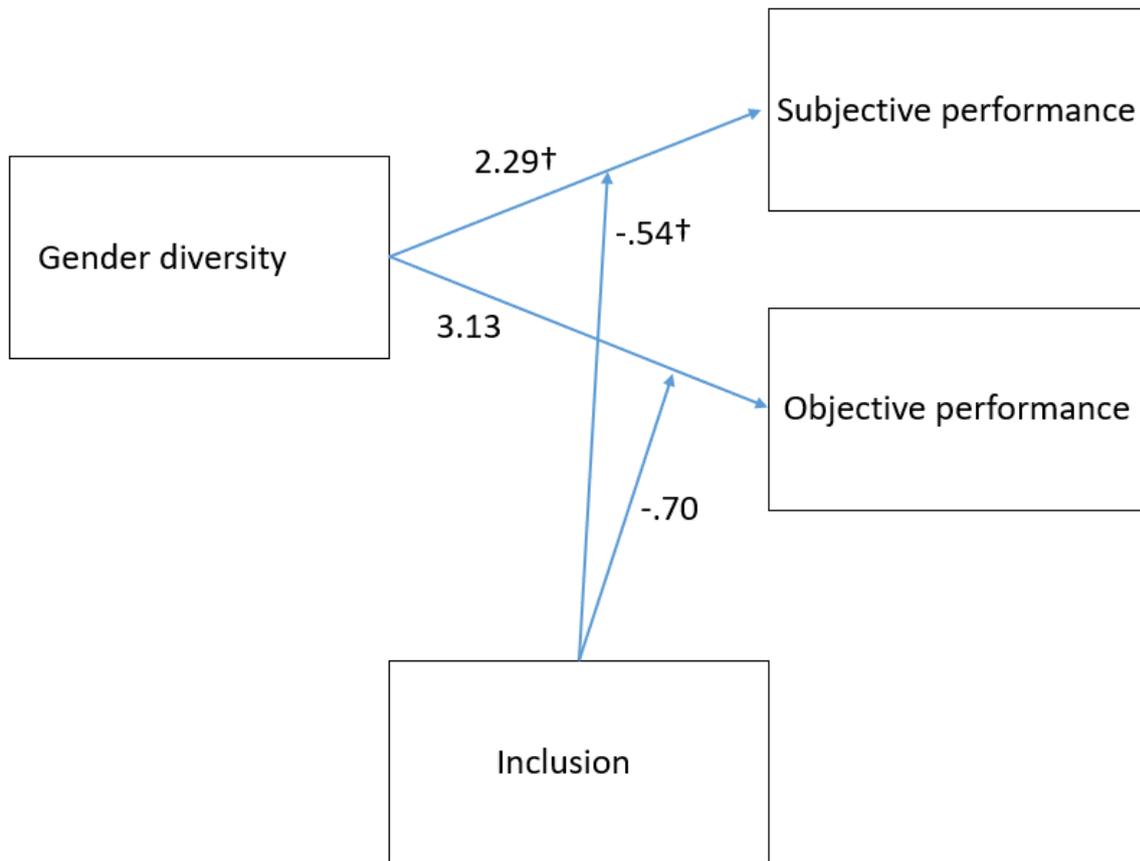
Note. † *p* <.10

Figure 4

Simple slope analysis
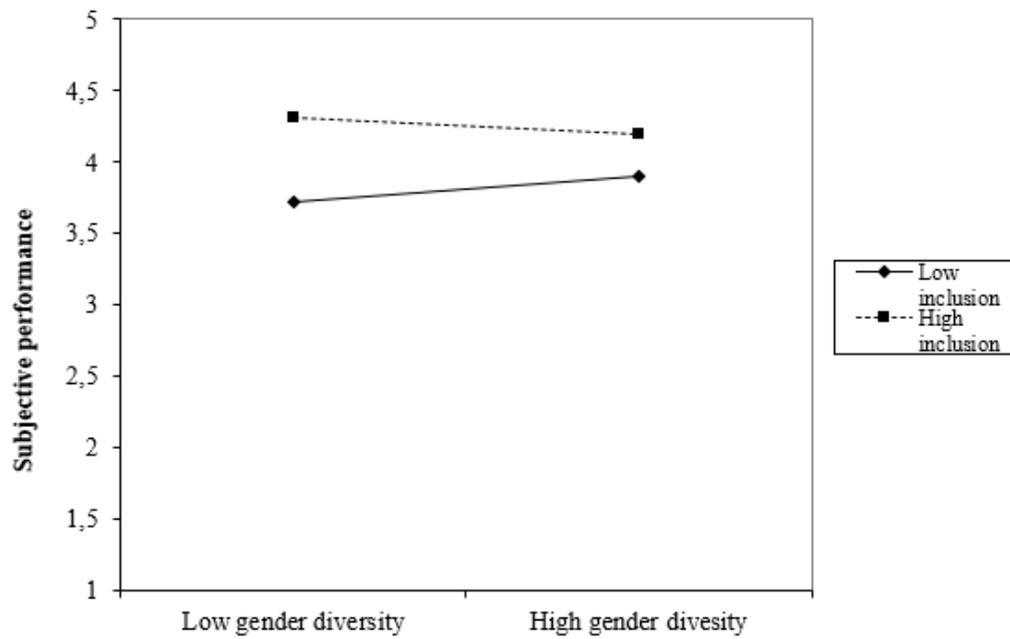
Figure 5

Simple slope analysis
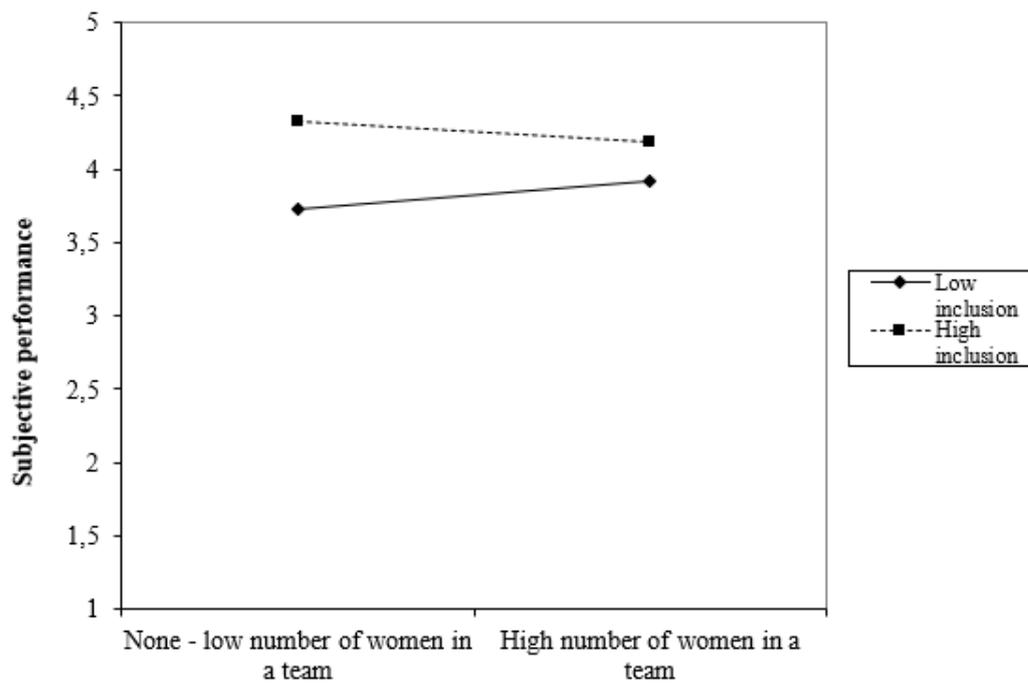
Figure 6

Moderation coefficients for nationality diversity – performance relationship with inclusion as moderator
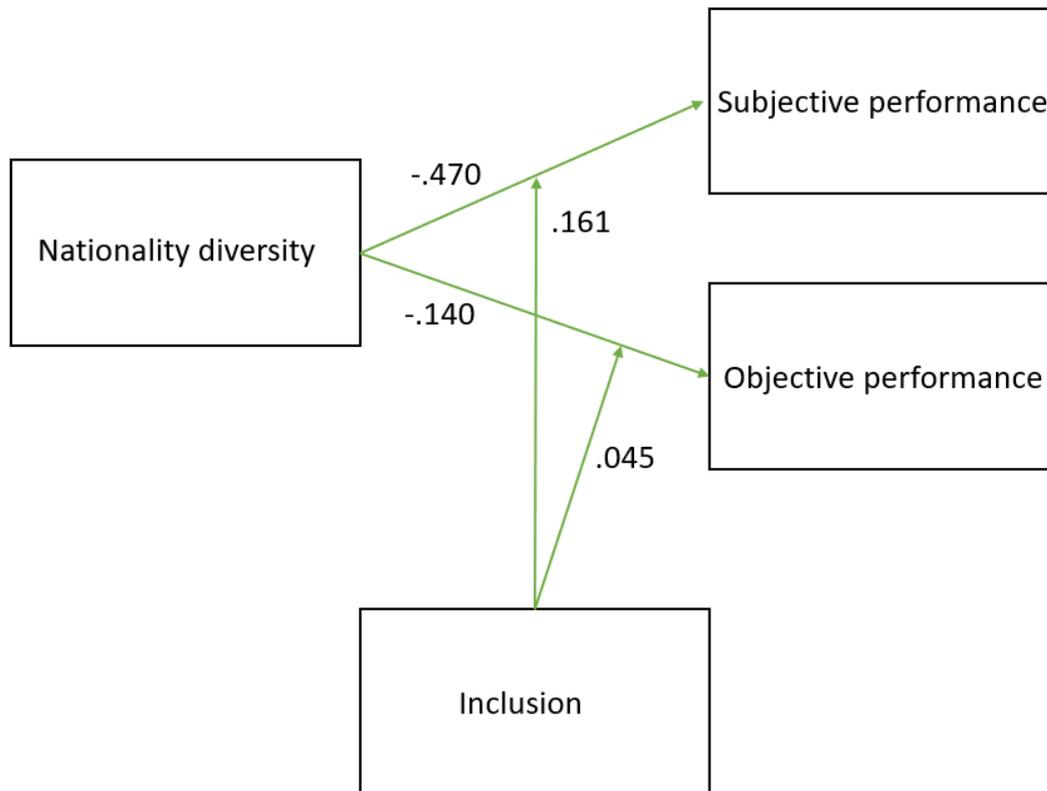
Note. No significant effects were found. † *p* <.10

Exploratory Analysis

Even though it was not hypothesized that the *deep-level diversity – performance* link would be moderated by *inclusion*, it was decided to perform this moderation analysis since no significant effect was found in the regression analysis. The moderation analyses of hypothesis 2 was repeated, but with *professional diversity* as independent variable, see Figure 7. The main effect of *professional diversity* on *subjective performance* shows no significant relationships (*p* = .73). The main effect of *performance* and *inclusion*, also shows no significant results for *subjective performance* (*p* = .20). Lastly, the moderation analysis was not significant either (*p* =.80).

Then, the main effect of *professional identity diversity* on *objective performance* was tested and the moderating effect of *team inclusion* on this relationship was analysed. The relationship between *professional diversity* and *objective performance* was not significant (*p* = .97). Also, no significant relation was found between *inclusion* and *objective performance* (*p* =

.77). And no significant moderation effect was found ($p = .97$). Therefore, no conclusions can be drawn about this model.

Since diverse teams in this study are mostly teams with more women an independent t-test was performed to investigate whether women rate their team *performance* higher than men. However, it was not found that women rated their performance higher than men, since no significant relation was found ($p = .42$) [3].

Figure 7

Moderation coefficients for the diversity – performance relationship with inclusion as moderator
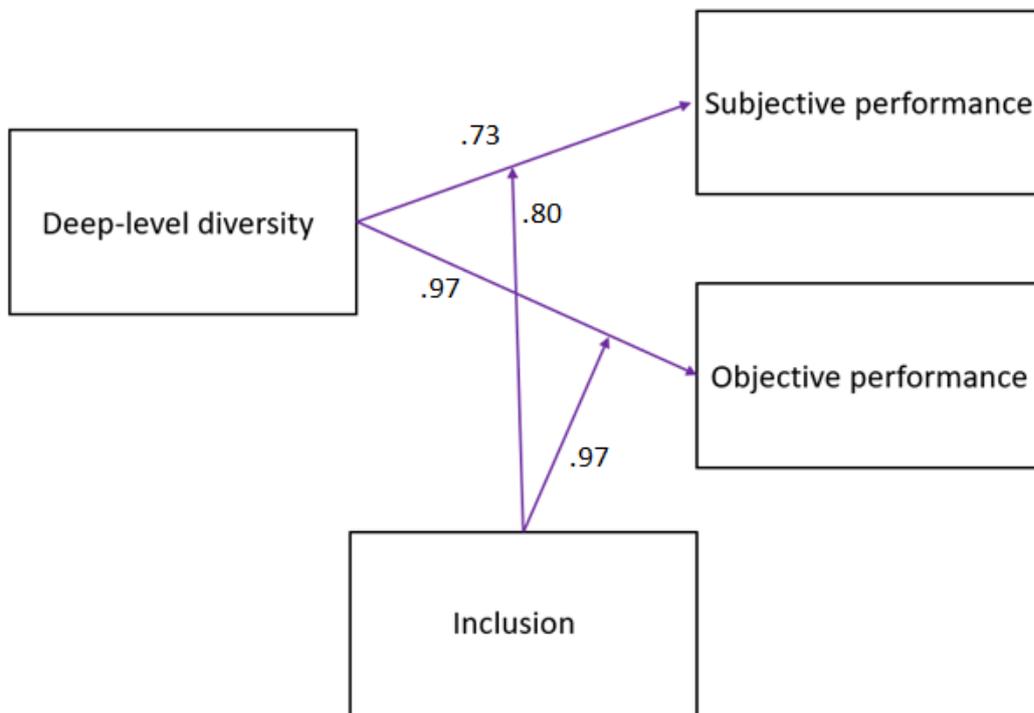
Note. No significant effects were found.

---

[3] To gain more insight in the *gender diversity – performance* relationship, two separate linear regression analyses were done. First with *gender diversity* as independent variable and *subjective performance* as dependent variable. No significant relationship was found ($t = .28$, $p = .78$). Second with the number of women as independent variable and *subjective performance* as dependent variable. Also, no significant relationship was found ($t = .51$, $p = .62$).

Discussion

This research was done to learn more about the *diversity – performance* relationship, since literature does not show consistent findings on this topic. The categorization-elaboration model (CEM) was used as a base for this study. It states on the one hand that when team members share their knowledge, this should result in better *performance* (information elaboration perspective). On the other hand, it is said that differences between team members can lead to ingroups and outgroups which is detrimental for *team performance* (social categorization process). So, based on CEM, it was likely that more *deep-level diversity* in a team would lead to better *performance* (information elaboration). However, this expectation could not be confirmed by the results of this study, so no proof was found for the information elaboration perspective of the CEM. It is possible that team members were not (sufficiently) aware of the professional profiles the other members. Which could be caused by not having spend enough time and effort on the exercise, since teams preferred starting immediately on the project rather than investing time in the exercise. This lack of time spend on the exercise could have led to not knowing each other well enough to benefit fully from everyone's expertise. On the other hand, Triana et al. (2021) found that different types of *deep-level diversity* (i.e., personality, values, cultural) have different outcomes. Diversity in values results in fewer positive team emergent states and team processes. So, it could be that the different types of *deep-level diversity* in this study (i.e., interest, values, competencies, and characteristics) changed the outcome of *performance*.

Secondly, it was expected that *surface-level diversity* would lead to worse *performance* (social categorization). However, when team members would *feel included*, it was expected that *surface-level diversity* would lead to better *performance*, but not when they did not *feel included*. In contrast, the results showed that when team members felt more included, *gender diversity* did not seem to increase *subjective performance*. When team members felt less or not included, more *gender diversity* led to higher *subjective performance*. These findings do not support the social categorization perspective from the CEM. It is difficult to clarify the moderating effect of *inclusion* on the *gender diversity – performance* link since it is unexpected and inconsistent with findings in literature (e.g., Joshi & Roh, 2009; Meadows & Sekaquaptewa, 2013). Though, positive effects of *gender diversity* on *performance* have also been found (Tshetshema & Chan, 2020). Alternatively, it could be questioned if men and women in the teams estimated their *performance* differently. If women estimated the *team performance* higher than men, than this could give a seemingly false result that *gender diverse*

teams rated their *performance* higher. However, the moderation effect was only marginally significant and the sample size was not large, so it is difficult to draw conclusions with certainty based on these results.

Third, it was predicted that more *nationality diversity* in teams would lead to lower *performance*. Since *nationality* is quite a salient characteristic, it could lead to ingroups and outgroups in a team which is disadvantageous for *team performance*. Yet, the results indicated that *nationality diversity* in teams did not predict *subjective performance*. Additionally, it was found that *inclusion* did not moderate the *nationality diversity – performance* relationship. So again, these results do not support the social categorization perspective of the CEM. Having nationality *diversity* in a team did not influence how team members rated their *performance*. It can be argued that the *nationality* differences in a team were not as big an issue as expected. Possibly since German and Dutch nationalities do not differ too much from each other. Alternatively, there is the possibility is that *inclusion* compensated the negative effect of *nationality diversity* which led to the insignificant relationship with *performance*.

Fourth, it was found that both *gender* and *nationality diversity* did not predict *objective performance*. Also, *inclusion* did not moderate these relationships. For gender, it could be that team members are used to work in *gender diverse* teams or do not have the prejudice that women perform less than men in technical teams. Vooren, Haelermans, Groot, and Maassen van den Brink (2022) and O'Dea, Lagisz, Jennions, and Nakagawa (2018) found that female students perform equally well as men in STEM education. This could mean that the preconceptions that men are smarter and perform better are not equally represented in schools as they are in businesses. This explanation would mean less social categorization in teams, which is beneficial for the *team performance*. For nationality, a similar reasoning as for *subjective performance* could be used. Since the social categorization process did not seem have an effect on *objective performance*, it could be argued that team members did not split into ingroups and outgroups or that the relationship was moderated by *inclusion*.

Lastly, while *feeling included* seems to predict the individual feelings of performance (*subjective performance*) it does not translate to the actual performance (*objective performance*) of the team. It could be that the duration of the project was too short (only one week). The support that was found for the moderating effect of *inclusion* (Bodla et al., 2018; Harrison et al., 2002; Nishii, 2013) came from studies in which team members worked together extensively and for a long(er) period of time than the current study. In this study, team members did work together everyday which would lead to feeling included in your team which could lead to the feeling that your team is performing well. That would explain the *inclusion – subjective*

performance relationship. However, working together for only one week is probably not enough to actually benefit from each others' skills and knowledge and make interactions go smoothly yet. So, the duration of the project week might have been too short for *inclusion* to increase the *objective performance*. On the other hand, this was a correlational study, which means that a causal conclusion cannot be drawn. Therefore, it could be that the relationship is reversed so that the individual feelings of performance (*subjective performance*) lead to better *inclusion*. Since the results (*objective performance*) are not known during teamwork, they cannot have an effect on *inclusion*, which could explain the lack of effect for *objective performance*.

Theoretical and Practical Implications

This study supports existing literature that shows the positive effects of *inclusion* on teamwork (Bodla et al., 2018; Harrison et al., 2002; Nishii, 2013). The results showed a significant relationship between *inclusion* and *subjective performance* and it confirmed the value of *feeling included* in a *gender diverse team* for *performance*. This suggests that schools and organizations can enhance *performance* by increasing *feelings of inclusion* of their students and employees, especially when they work in *gender diverse teams*. It is expected that, though this research does not show a direct effect with *objective performance*, over time *subjective performance* will lead to better *objective performance* and other outcomes like, less turnover and more commitment (Nishii, 2013).

Additionally, quite some insignificant and unexpected results have been found in this study. For example, the non-significant relationship between *deep-level diversity* and *performance* or the inconsistent findings of *gender diversity*. It could be questioned if the CEM applies for student teams. Maybe student teams are not diverse enough in knowledge and skills in comparison to work teams. Or the preconceptions (e.g., that women are not as valuable as men in technical teams) are not as strong in *gender diverse* student teams than in work teams.

Limitations and Suggestions for Further Research

This study took place in an actual project week with participants who had to do well on their project to pass. Even though the environment was realistic, there were some limitations in this study. Firstly, the sample size, the data of only 71 teams were used in the analysis. Some relationships were marginally significant. So, research with a larger sample size could draw conclusions with more certainty. Replication of the current results would be a verification of the conclusions that were drawn in this study.

Secondly, the composition of the teams was a limitation, for both gender and nationality. In the current study, gender diversity in teams consisted mostly of many men and a minority of women. This might be a truthful representation in technical team, but gender diverse teams consisting of a combination of the majority being men or women leads to more generalizable and insightful results to whether gender diversity is related to the variables in this study. Similarly to gender, nationality diversity in teams in this study consisted mostly of many Dutch and a minority of German members. Additionally, Dutch and German cultures are not very different from each other. Therefore, it would be insightful to study nationality diverse teams that include more nationalities and include teams with a variety of the majority being Dutch or other nationalities which would lead to more general results and thus conclusions.

Another limitation of this study was the duration. The project was only one week, in this week, the project had to be finished. Therefore, the time that team members wanted to use to work on their projects now partly had to be invested in the exercise. This led to complaints about the exercise and its duration. A suggestion for further research would be to shorten the exercise to make it less of a distraction from the actual project the teams should be working on. Additionally, guidance and support might increase the efficiency of the exercise. Moreover, the number of measurements that were done during this research (each variable (i.e., *performance* and *inclusion*) was measured once), could be increased. In future research a longitudinal study could get a more complete insight in the variation of those variables in team members over time.

Lastly, it would be interesting to setup an experimental study to learn more about the *inclusion – performance* relationship. The intervention would be an activity to build a team mentality before the teams start working on the actual project to increase *feelings of inclusion*. All teams would then have to fill out a survey where they indicate how *included* they personally feel as well as how *included* they think team members feel in general. The same questions should be answered at the end of the project to see if these *feelings of inclusion* changed during the teamwork. The teams in the control group should either spend the same amount of time together before starting their projects or 'time spent together before the project' should be controlled for in the analysis. Additionally, it would be interesting to study different durations of projects. It is possible that the effect of the activity decreases when teams work together for a longer time.

Conclusion

This study tried to answer the questions when *diversity* affects *performance*. No significant relationship was found between *deep-level diversity* and *performance*. For *surface-level diversity*, the findings were mixed. It was found that *feeling included* in a *gender diverse* team is beneficial for the quality of *performance*. However, no significant relationships were found regarding *nationality diversity* and *performance*. Lastly, a significant relationship between *inclusion* and *subjective performance* was found. To summarize, this research found that *team inclusion* is important for good *performance*, especially in *gender diverse* teams.

References

Blau, P. M. (1977). *Inequiality and heterogeneity*. Free Press.

Bodla, A. A., Tang, N., Jiang, W., & Tian, L. (2018). Diversity and creativity in cross-national teams: The role of team knowledge sharing and inclusive climate. *Journal of Management and Organization*, *24*(5), 711–729. https://doi.org/10.1017/jmo.2016.34

Booy, C., Jansen, E. J. M., Joukes, G. W. M., & Van Schaik, E. J. (2012). *Trend analysis gender in higher STEM education*. VHTO. https://www.vhto.nl/fileadmin/user_upload/documents/publicaties/Internationaal/LR_V HTO_GENDER_ENGELS_TOTAAL.pdf

Dawson, J. (n.d.). *2-way_linear_interactions.xls*. Retrieved November 9, 2021, from http://www.jeremydawson.co.uk/2-way_linear_interactions.xls

De Dreu, C. K. W., & Weingart, L. R. (2003). Task versus relationship conflict, team performance, and team member satisfaction: A meta-analysis. *Journal of Applied Psychology*, *88*(4), 741–749. https://doi.org/10.1037/0021-9010.88.4.741

Endedijk, M., Ten Berg, D., Nieuwenhuis, M., Van Veelen, R., & Van Hattum, N. (2019). *Who is the technical professional ? A description of five identity profiles* (p. 16).

Hackman, J. R. (1987). The design of work teams. In J. Lorsch (Ed.), *Handbook of organizational behavior* (pp. 315–342). Englewood Cliffs, NJ: Prentice Hall.

Harrison, D. A., Price, K. H., & Bell, M. P. (1998). Beyond Relational Demography : Time and the Effects of Surface- and Deep-Level Diversity on Work Group Cohesion Author (s): David A . Harrison , Kenneth H . Price and Myrtle P . Bell Reviewed work (s): Source : The Academy of Management Journal , Vol. *The Academy of Management Journal*, *41*(1), 96–107.

Harrison, D. A., Price, K. H., Gavin, J. H., & Florey, A. T. (2002). Time, Teams, and Task Performance: Changing Effects of Surface- and Deep-level Diversity on Group Functioning. *Academy of Management Journal*, *45*(5), 1029–1045. https://doi.org/https://doi.org/10.5465/3069328

Hayes, A. F. (2013). *Introduction to Mediation, Moderation, and Conditional Process Analysis: A Regression-Based Approach*. Guilford Press.

Horwitz, S. K., & Horwitz, I. B. (2007). The effects of team diversity on team outcomes: A meta-analytic review of team demography. *Journal of Management*, *33*(6), 987–1015. https://doi.org/10.1177/0149206307308587

Hui, J. S., & Farnham, S. D. (2016). Designing for Inclusion: Supporting Gender Diversity in

Independent Innovation Teams. *Proceedings of the 19th International Conference on Supporting Group Work (GROUP '16). Association for Computing Machinery*, 71–85. https://doi.org/10.1145/2957276.2957290

Jackson, S. E., & Joshi, A. (2011). Work Team Diversity. In S. Zedeck (Ed.), *APA Handbook of Industrial and Organisational Psychology* (Vol. 1, pp. 651–686).

Jansen, W. S., Otten, S., van der Zee, K. I., & Jans, L. (2014). Inclusion: Conceptualization and Measurement. *European Journal of Social Psychology*, *44*(4), 370–385.

Joshi, A., & Roh, H. (2009). The role of context in work team diversity research: a meta-analytic review. *Academy of Management Journal*, *52*(3), 599–627. https://doi.org/10.5465/amj.2009.41331491

Koo, T. K., & Li, M. Y. (2016). A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *Journal of Chiropractic Medicine*, *15*(2), 155–163. https://doi.org/10.1016/j.jcm.2016.02.012

Larkey, L. K. (1996). Toward a Theory of Communicative Interactions in Culturally Diverse Workgroups. *The Academy of Management Review*, *21*(2), 463–491. https://doi.org/10.2307/258669

Meadows, L. A., & Sekaquaptewa, D. (2013). The influence of gender stereotypes on role adoption in student teams. *120th ASEE Annual Conference & Exposition*. https://www.researchgate.net/publication/287234332_The_influence_of_gender_stereotypes_on_role_adoption_in_student_teams

Möwes, R. A. (2016). *The career compass: Discovering diversity in STEM students' professional identity and its effects on their intended career choice*. http://essay.utwente.nl/69306/

Nishii, L. H. (2013). The benefits of climate for inclusion for gender-diverse groups. *Academy of Management Journal*, *56*(6), 1754–1774. https://doi.org/10.5465/amj.2009.0823

O'Dea, R. E., Lagisz, M., Jennions, M. D., & Nakagawa, S. (2018). Gender differences in individual variation in academic grades fail to fit expected patterns for STEM. *Nature Communications*, *9*. https://doi.org/10.1038/S41467-018-06292-0

Shore, L. M., Randel, A. E., Chung, B. G., Dean, M. A., Ehrhart, K. H., & Singh, G. (2011). Inclusion and diversity in work groups: A review and model for future research. *Journal of Management*, *37*(4), 1262–1289. https://doi.org/10.1177/0149206310385943

Triana, M. del C., Kim, K., Byun, S. Y., Delgado, D. M., & Arthur, W. (2021). The Relationship Between Team Deep-Level Diversity and Team Performance: A Meta-Analysis of the Main Effect, Moderators, and Mediating Mechanisms. *Journal of Management Studies*,

58(8), 2137–2179. https://doi.org/10.1111/JOMS.12670

Tshetshema, C. T., & Chan, K. Y. (2020). A systematic literature review of the relationship between demographic diversity and innovation performance at team-level. *Technology Analysis & Strategic Management*, *32*(8), 955–967. https://doi.org/10.1080/09537325.2020.1730783

Tsui, A. S., Egan, T. D., & O'Reilly, C. A. (1992). Being Different: Relational Demography and Organizational Attachment. *Administrative Science Quarterly*, *37*, 549–579. https://doi.org/10.5465/ambpp.1991.4976867

Van Knippenberg, D., De Dreu, C. K. W., & Homan, A. C. (2004). Work group diversity and group performance: An integrative model and research agenda. *Journal of Applied Psychology*, *89*(6), 1008–1022. https://doi.org/10.1037/0021-9010.89.6.1008

Van Knippenberg, D., Van Ginkel, W. P., & Homan, A. C. (2013). Diversity mindsets and the performance of diverse teams. *Organizational Behavior and Human Decision Processes*, *121*(2), 183–193. https://doi.org/10.1016/j.obhdp.2013.03.003

Van Veelen, R., & Ufkes, E. G. (2019). Teaming Up or Down? A Multisource Study on the Role of Team Identification and Learning in the Team Diversity–Performance Link. In *Group and Organization Management* (Vol. 44, Issue 1). https://doi.org/10.1177/1059601117750532

Vooren, M., Haelermans, C., Groot, W., & Maassen van den Brink, H. (2022). Comparing success of female students to their male counterparts in the STEM fields: an empirical analysis from enrollment until graduation using longitudinal register data. *International Journal of STEM Education*, *9*(1), 1–17. https://doi.org/10.1186/S40594-021-00318-8/TABLES/8

Appendix A

Questions used for Analysis from Questionnaire 1

Q7 What is your age?

Q8 What is your gender?

Q9 What is your nationality?

Q11 What is the name of your educational institution?

Q12 What is the name of your educational study program?

Q13 Which study year are you currently in?

Q24 Please fill in the scores with the test results from the **Professional Profile Test** (Excel sheet)   below,

All-rounder _____

Analyst _____

Individualist _____

Team player _____

Innovator _____

Appendix B

Questions used for Analysis from Questionnaire 2

Q22 Howe many team members did you know prior to the project week? (Select "0" if you did not know anyone before)

Q16 What are your **reflections** about how the team members are included in the project assignment this week?   *In this team...*

… I think that all team members feel included.

… I think everyone is accepted for who they are.

… I think everyone makes a unique contribution.

… I think that differences between team members are valued.

>*5-point Likert scale, i.e., Strongly disagree – Strongly agree*

Q11 In your opinion how well does your project **team perform**? Please grade your project team's performance on...

… accomplishing project goals.

… the quality of the project work.

… the quantity of the project work.

… the overall performance level.

>*5-point Likert-type scale, i.e., Terrible - Excellent*

Appendix C

Development Exercise and Material

The exercise consisted of an individual and team part. First, all team leaders received an introduction (see Figure 8) and instruction sheets (see Figure 9) where all the necessary information to complete the exercise was summarized. All team members started with individually filling in the short version of the Career Compass (Appendix E) that was developed for this study. The result were profile scores for all five profiles. Those scores were the starting point for the individual worksheet (see Figure 10) on which the scores had to be written down as well as a top 3 based on those scores. Then the profile boards (see Figure 11) were distributed with detailed information about the profiles. With this information the profile that fits someone best could be chosen. Then there was space to write down 3-5 characteristics of this profile that the individual felt most fitting for him/her. This was summarized in a personal profile card. Now the individual worksheet was completed.

For the team part, the team leader would start and guide the conversation and help questions were given as an example to stimulate the conversation. Firstly, all members introduced themselves to each other based on the profile and characteristics they chose and the skills they have. Secondly, those personal profile cards were combined at the team board (see Figure 12). Additionally, the team board had tasks written on them that could be divided among team members with pink, yellow, and blue post-its. If someone wanted to do a task or thought (s)he was good at it, it would be blue. If someone wanted to learn to do a task or was not sure about his/her ability, it would get a yellow post-it. And if someone did not want to do a certain task, it would get a pink post-it. The team leader would guide and oversee this process as well. The clarity of the instructions was not tested before it was used by the project teams.

Figure 8

Introduction for team leader

LEARN MORE ABOUT EACH OTHER!
START THIS RIGHT AFTER COMPANY PRESENTATION

Before you start working on your company assignment, you will start working on this interactive task to find out how everyone can contribute to the team tasks. Why?

People differ: not only in terms of your study background, but also in terms of what you find important, your interests, your personality and your competences. Together this forms your professional profile. If you are familiar with your own professional profile and know about the professional profile of your team members, you will be able to make optimal use of each other's talents, collaborate well, have fun and come up with good and innovative solutions.

The task consists of three parts:

1. Learn more about yourself: all team members answer questions online and will find out about their own individual profile.

2. Learn more about your teammates: you will share your profiles and find out about the unique characteristics of your team and what you find important to work on as a group.

3. Prepare for the task: you will find out how the characteristics of everybody can be used during the project week.

You will need about 45 minutes for this task and then you are perfectly prepared for the project week!

Do you need assistance? Ask the teaching assistant for help or send us a Whatsapp message to: +316 28 32 41 84 with your room and group number or come and find us at the Project Management Office (W2.39). We will come to help you as soon as possible!

Participation in this task is a mandatory element of the project week. Saxion collaborated with the University of Twente to design and evaluate this interactive task. The outcomes will be used to further improve the project week for next years.

Please use the material enclosed in this folder to start the task

Figure 9

Instructions for team leader

LEARN MORE ABOUT EACH OTHER!
INSTRUCTIONS TEAM LEADER

START HERE!

15 min.

STEP 1: Learn more about yourself

You need this:

1. Individual worksheet (one per member)
2. 2 Profile boards (do not distribute yet)
3. Pen/pencil
4. Laptop

To do:

1. Distribute the individual worksheets and put the profile boards upside down on the table.

2. Let everyone work on the individual worksheet, including yourself!

3. Check if everybody has completed the personal profile card. Some people might by quicker than others, so be patient and give everybody the time needed to complete.

STEP 2: Learn more about your teammates

15 min.

You need this:

1. Team board
2. Pen/pencil

To do:

1. Write down the names of your teammates on the board.

2. Put the team number (e.g., DEMCON01) on the team board

3. Start left: ask every team member:

- What profile did you select for yourself and why?

- Did you change the original and why?

- Which key characteristics – related to the profile – do you want to put on the team board?

4. Write down the key characteristics they want to put on the team board.

5. Count which profiles are most present. Write down the number in the upper right corner. Now discuss the following questions:

- What profiles over- or underrepresented?

- What does this mean for your project?

6. Discuss the characteristics that everybody wrote down:

- What are the similarities?

- What are unique characteristics?

- Ask everybody to underline their most important characteristic on the board.

LEARN MORE ABOUT EACH OTHER!
INSTRUCTIONS TEAM LEADER

STEP 3: Prepare for the task

15 min.

You need this:

1. Team board
2. Pen/pencil
3. Post-its

To do:

1. Give each team member post-its.

2. Invite all team members to think about each task and how that fits them. They can indicate this by using post-its. You can also leave tasks empty if you do not have a clear preference.

> Blue = this fits me! Write down one or two words what aspect of this task fits you.
>
> Challenge = this task is something I would like to be engaged in, but also want to learn from. Write down one or two words about what you want to learn.
>
> Pink = this does not fit me.

For example: if person X feels like designing fits their profile, than he/she should put a green post-it in their column and at the row of the report that indicates "designing lay-out".

3. Discuss the following questions based on the outcomes:

- Are their more or less popular tasks?
- How can team members help each other?

4. Fill in the last row of the table. Where are team members focusing on? What is their role during the project week?

One of the teaching assistants will come and take picture of your team board after you finished. This will serve as a proof that you completed this task. You can keep this board with you during the project week and make changes or add things if you want.

Enjoy the project week!

This task fits me!

This task does not fit me

This task feels like a challenge

Figure 10

Individual worksheet

STEP 1. INDIVIDUAL WORKSHEET

1. Fill in this URL on your laptop:
https://bit.ly/2RUGdmC

2. Finish the survey. Do not forget to mention the results from the test in the online survey.

3. What are your results from the test? ——————————→

4. According to the test, the profile with the highest score fits you best.

Write down here your top 3: **1.** ...   **2.** ...   **3.** ...

My scores from the test

☐	**All-rounder** 🔴
☐	**Analyst** 🟢
☐	**Individualist** 🔵
☐	**Team player** 🟡
☐	**Innovator** 🟣

5. Take one profile board from the pile. Read the profile description of the profile of your highest scoring profile. Underline the aspects that fit you well.
? Do you have the feeling that it does not fit you well? Read then the description of your number 2 (or if needed also 3). Make a decision which profile fits you best and indicate this on your personal profile card below.
Are there some words unfamiliar to you? For each group there is a Dutch translation available.

6. Now check the characteristics that you underlined and think of the project week and the things you are going to do. Pick 3-5 of these characteristics that you think are relevant for the project week. Write these down on your personal profile card.

7. The next step will be that you will be asked to introduce yourself, the profile you selected and your 3-5 characteristics to the rest of your team. But first wait till the rest of the team is finished with this part.

Personal Profile Card

Name:

This profile fits best

🔴 🟢 🔵 🟡 🟣

My 5 characteristics:
1.
2.
3.
4.
5.

Figure 11

Profile board

PROFILE BOARD

People in the 'Innovator' profile typically enjoy **challenging** and **intellectually stimulating work**. They value having **good career perspectives**, which is likely to be more important than family relationships and adhering to traditions. Typically, they do not need predictability, preferring instead to **experiment** and seek out **challenges**. Volunteering and societal themes such as well-being and the natural environment are often of little importance to innovators.

Innovators are full of **initiative, creative in their thinking** and very **open to new ideas**. They tend to appear **optimistic, self-assured** and **assertive**. When collaborating with others, they may sometimes run ahead of the group in their enthusiasm. Therefore, ensuring to complete tasks and initiatives before starting new ones and keep everyone involved can be a learning point.

Innovators have a lot of **confidence** in their **competencies** when it comes to **entrepreneurship, collaboration** and **analyzing problems**. Designing systems or products and making the most of changes in the environments are also competencies that come easy to them. They tend to struggle to structure their work and organize their time efficiently.

In short, innovators like to understand how things work and are often interested in creative, artistic and innovative activities. They prefer to minimize screen time, spending their time working with their hands instead.

People in the 'Team-player' profile value **social etiquette** and having good **family relationships**. These values also tend to surface in their community spirit, interest in local communities and a desire to maintain a good **work-life balance**. Team players care about being in good health, they value **certainty** and appreciate a **comfortable life**. They are generally less interested in intellectual stimulation.

Team-players tend to be **socially pro-active**; they seek contact and touchpoints with others and during meetings they are active participants. They often have a **grounded and pragmatic view** on matters and are focused on **action** rather than analysis. Their outgoing nature helps to make connections between people when collaborating, however team-players may at times be too attention-seeking and benefit from allowing others more time to speak. While team-players like to take action, they may at times benefit from being more open to reflecting on creative ideas. Team-players tend to be confident in their competencies regarding **managing other people**, and team **collaboration**. They evaluate their own analytical and research skills more negatively though. Also, designing systems or products and working in intercultural settings might me more difficult to them.

In short, team-players are interested in social interaction and the local communities they are part of. They prefer to act and get things done rather than analyze and think things through, which surfaces in their pragmatic approach.

People in the 'All-rounder' profile have many different motivational drivers. They value **intellectual challenges** and are keen to support other people and **contribute to society**. All-rounders like **to enjoy life**. They generally care about having good **family relationships**, their **health** and **career**, while to a lesser degree they also value predictability.

Many all-rounders seem curious and have a wide **variety of interests**. They tend to be open to **new ideas** and participate actively in conversations about these. Many all-rounders are well organized in their approach to work, which may help to satisfy their diverse motivational interests.

All-rounders have great confidence in their competencies. In particular skills such as **management, collaboration**, dealing with change and working in **international environments** comes easy to them. They are also positive about their analytical abilities, their ability to design systems or products, to mentor others and to organize their own time for maximum effectiveness.

In short, all-rounders have broad interests encompassing people and society, collaboration with others and creativity. They tend to be curious and enjoy practical activities. This diversity of interests could make it difficult for some all-rounders to make decisions about what goals to pursue.

People in the 'Analyst' profile love **intellectual stimulation**. They are driven to **understand problems** and **find solutions**, in particular when this benefits other people and society. Analysts are often satisfied having a **modest** lifestyle without unnecessary luxuries. Most analysist are not interested in status or power; for them career success revolves around developing their **expertise** and solving increasingly **complex analytical issues**.

Most analysts appear **independent** and **introverted**; they tend to **listen** more than they talk. At times, this makes it hard for others to understand what thoughts analysts are having on their mind.

Analysts typically make a **friendly** and **reliable** impression. They tend to be somewhat sensitive to work pressure or may doubt the quality of their own work.

Analysts are generally confident in their **analytical skills**. They tend to believe in their abilities to conceptualize high quality research, execute the studies and write up the results. Analysts are significantly less confident in their competencies regarding management, collaboration, flexibility in times of change and the creative design of systems or products.

In short, analysts enjoy working independently on complex analytical issues. In their limited interaction with others, they are focused on the contents of their work and tasks at hand.

People in the 'Individualist' profile value a **comfortable** and **simple life**; they are not looking for challenges for their own sake. Individualists are most comfortable on their own and typically do not need much contact, interaction or activities with other people. **Gaming** is a favorite way to spend their free time and some individualists also enjoy **working with their hands**. Individualists tend to be less interested in topics concerning society.

In their work behavior, individualists may seem somewhat **unorganized**; they may wait for deadlines to come really close and work with bursts of action. When they are amongst others, they tend to **listen** rather than speak. When working on projects, it can be a challenge for individualists to stay actively involved with the team. They enjoy working on **practical matters**, appear **task-focused** and prefer for others to take the lead. Often, they make a **calm** impression on others, which can however also make them appear indifferent.

Individualist are quite **critical** when evaluating their own competencies. In particular when these competencies are related to interaction with other people, for example managing, collaborating with or mentoring others. Because many individualists prefer to be on their own and not pro-actively seek out challenges, they may not have practiced these competencies as often.

In short, many individualists like to work independently on clearly defined tasks. They tend to feel most comfortable with technical or functional tasks they can complete on their own.

Figure 12

Team board

TEAM   …

… X   … X   … X   … X   … X

My 5 characteristics: 1. 2. 3. 4. 5.	My 5 characteristics: 1. 2. 3. 4. 5.	My 5 characteristics: 1. 2. 3. 4. 5.	My 5 characteristics: 1. 2. 3. 4. 5.	My 5 characteristics: 1. 2. 3. 4. 5.	My 5 characteristics: 1. 2. 3. 4. 5.	My 5 characteristics: 1. 2. 3. 4. 5.	My 5 characteristics: 1. 2. 3. 4. 5.	**TASKS**
								Background research on problem / requirements
								Come up with innovative ideas and solutions
								Reporting findings on paper
								Presenting the findings & communication with client
								Designing the final products (solution / poster / report)
								Project management and team support
								WHAT IS YOUR ROLE DURING THE PROJECTWEEK?

Appendix D

Development Shortened Version Career Compass

The Career Compass (Möwes, 2016) is a tool that helps students gain insight in who they could be as a technical professional in their career. This tool has approximately 100 items that need to be answered to receive a profile. Since in this case it was only a part of the exercise, it needed to be shorter to make sure that the participants did not have to spend too much time working on the exercise. The Career Compass has items four categories: interests, competencies, values, and personality. The statements for these categories are divided in 29 factors, for example, the factor 'social measures' is a part of the category 'interests'. To reduce the number of items for the intervention, descriptive analysis and the ANOVA test were used to find the factors that predicted the profiles the most. When the most predicting factors were found (reduced from 29 to 13), the two best predicting items for each of those factors were chosen. If one or both items were irrelevant for the participants, a more relevant item was chosen. So, two items for all factors resulted in a total of 26 items.

The statements were placed in an Excel-file (see Appendix E) that could be downloaded after the first survey was completed. The participants had to select 'below average', 'average', or 'above average' for each statement while comparing themselves with the average technical student. The answer 'below average' got a score of 1, 'average' got a score of 2, and 'above average' got a score of 3. Since every factor was measured by two items, all factors could get a score from 2 to 6. A factor can make a distinction between profiles that a participant could be if (s)he answered the item with below or above average. So, if the total score is 4, no profile gets a point. However, if the total score is 2 or 3, the profile(s) that is/are not predicted by the factor get(s) a '-1' and if the total score is 5 or 6, the profile(s) that is/are predicted by the factor get(s) a '+1'. In the end, all scores of the profiles are added (absolute value). Since not all profiles can get just as many points, the absolute value is divided by 26 (the maximum score that can be achieved by one of the profiles) and then multiplied by the maximum score of that specific profile. This gives the final score for each profile which is then shown to the participant so that (s)he knows which profile fits him/her best.

Respondents

The Excel sheet with statements was tested on functionality, clarity, and providing useful answers. The respondents were (study) friends and family, who were either studying or had studied at a University or a University of Applied Sciences. In total, there were eight

respondents, four male and four female. Five of them were studying at a University and three were studying/had studied at a University of Applied Sciences. See Table 9 for a summary of the information of the participants.

Table 9

Descriptive characteristics participants

	n	%
Gender		
Male	4	50
Female	4	50
Education		
University	5	62.5
University of Applied Sciences	3	37.5

Procedure

Three files were sent to the participants. The first one was an explanation on how to answer the statements, the second one was the Excel file with the 26 statements, and the third file was the explanation of all five profiles. The pilot test of those statements consisted of two parts. Firstly, the participants had to answer the 26 statements and secondly, they had to read the text about the five profiles and they had to write down their numbers 1, 2, and 5, with 1 best fitting and 5 worst fitting profile. To check the validity of this short version of the Career Compass, the outcome of the statements was compared to the ranking of the participants to see if it was similar.

Data analysis

The instruction text, Excel sheet with the statements, and the information text on the profiles were found understandable and clear. Only one spelling error was found and corrected. The responses that were received showed that the calculations behind the statements presented fitting profiles regarding the participants numbers 1, 2, and 5.

Appendix E

Career Compass (shortened)

Professional Profile Test	
INSTRUCTIONS: Think about the average technical student. How do you see yourself compared to the average technical student? Please take a moment to think about yourself and your own characteristics, skills, interests, and values. Then, select for each statement below if you see yourself as 'below average', 'average', or 'above average' compared to the average technical student. To select your answer click on the arrow in column B.	Compared to the average technical student, I score ...
01. I like meeting up with friends.	
02. I like strategic games.	
03. I like social activities.	
04. I like solving puzzles.	
05. I am good at leadership and management.	
06. I am good at conducting research.	
07. I am good at thinking analytically.	
08. I am good at dealing with uncertainty.	

09. I am good at team work.	
10. I am good at developing research ideas.	
11. I am good at adapting to new situations.	
12. I am good at collaborating with people outside my own study program.	
13. I am good at networking.	
14. I am good at developing solutions to complex problems.	
15. I find respect for tradition important.	
16. I find ethical responsibility important.	
17. I find lifelong learning important.	
18. I find routine and structure important.	
19. I find a just world important.	
20. I find politeness important.	
21. I find stability important.	
22. I find intellectual stimulation important.	
23. I am introverted.	
24. I am open-minded.	
25. I am reserved.	
26. I am imaginative.	

INSTRUCTION: Make sure you have answered all statements. Click on the tab 'Results' to get your test results for the Professional Profile Test!

If nothing was filled in, the following was visible:

Results Professional Profile test

There are 5 profiles of technical professionals. Below you find your test scores for each of the 5 profiles.

Professional profiles

All-rounder

Analyst

Individualist

Team player

Innovator

Please note:

1. You only get your profile scores if you answered all statements.

2. If all your scores on the profiles are zero, this means we cannot compute your profile scores.

Please read the statements again and make sure you do not score 'average' on all statements.

If all statements were answered, the following was visible (based on the answers that were given)

Results Professional Profile test

There are 5 profiles of technical professionals. Below you find your test scores for each of the 5 profiles.

Professional profiles	Your test scores The profile with the highest score fits you best.
All-rounder	8
Analyst	3
Individualist	8
Team player	9
Innovator	12

Please note:

1. You only get your profile scores if you answered all statements.

2. If all your scores on the profiles are zero, this means we cannot compute your profile scores. Please

read the statements again and make sure you do not score 'average' on all statements.

INSTRUCTIONS: PLEASE FOLLOW THE STEPS BELOW ONCE YOU COMPLETED THE TEST

1. Go back to the online questionnaire and insert your test scores for each of the 5 profiles.

2. Also insert your test scores on your individual worksheet (STEP 1: INDIVIDUAL WORKSHEET, point 3) and follow the instructions.