

Jacks-of-all-trades? The effect of balanced skills on team performance*

Laura Rosendahl Huber[†], Randolph Sloof[‡] and Mirjam Van Praag[§]

Abstract

Previous empirical studies have shown that solo entrepreneurs benefit from having balanced skills: Jacks-of-All-Trades (JATs) are better entrepreneurs than specialists are. Nowadays however, the majority of entrepreneurs start up and run ventures together in teams. In this paper we test whether the effect of more balanced skills is also positive in a team of entrepreneurs. We also explore whether (a lack of) individual balanced skills can be substituted by combining the skills of various specialists within one team. Our field experiment studies teams of children participating in an entrepreneurship education program. Based on pupils' precisely measured level of verbal and mathematical ability, we exogenously compose 179 teams separated into four different types: JAT teams, math-specialist teams, verbal-specialist teams and mixed specialist teams. Our results show that balanced skills are beneficial to team performance, and that it is hard to substitute individual balanced skills by combining different specialists within one team.

JEL-codes: C93, D83, J24, L25, L26, M13

Keywords: skill balance, team diversity, team performance, entrepreneurship, field experiment

*This research was supported by the Amsterdam Center for Entrepreneurship. We would like to thank seminar audiences at the SMS conference (Tel Aviv), EALE conference (Turin), Amsterdam Center for Entrepreneurship (University of Amsterdam and VU University) and Tinbergen Institute for helpful comments.

[†]Amsterdam School of Economics, University of Amsterdam, Valckenierstraat 65-67, 1018 XE Amsterdam, Netherlands. Email: l.rosendahlhuber@uva.nl

[‡]Amsterdam School of Economics, University of Amsterdam and Tinbergen Institute. Email: r.sloof@uva.nl

[§]Copenhagen Business School, Dept of Innovation and Organizational Economics, Kilevej 14a, 2000 Frederiksberg, Copenhagen, Denmark; Tinbergen Institute and IZA. Email:.mvp.ino@cbs.dk

1 Introduction

Nowadays, a substantial and growing share of businesses are started up and run by entrepreneurial teams instead of solo entrepreneurs.¹ The entrepreneur stereotyped as a 'lone wolf' has become less and less representative. Especially in knowledge intensive industries, such as the high-tech industry, the majority of start-ups is founded by teams (Lechler, 2001). Because successful entrepreneurs are vital for the economy and its growth and innovation, this immediately raises the question of effective team composition. In this paper we empirically explore this question focusing on the *skill* composition of successful entrepreneurial teams.

Following the reasoning of various (classic) economists, most notably Marshall, influential models have been developed assuming that ability is vital for individual entrepreneurship (Lucas, 1978; Jovanovic, 1982). Indeed, empirical evidence supports the assertion that ability is an important driver of entrepreneurial performance, see e.g., Gompers et al. (2010). In these studies, ability is often measured in terms of cognitive skill or intelligence (e.g., Hartog et al., 2010) or education (e.g., Parker and Van Praag, 2006; Gennaioli et al., 2013; Van Praag et al., 2013). One of the main conclusions from this literature is that the performance of entrepreneurs is determined by skill rather than luck.²

Despite the relatively solid knowledge of the role of skill for individual entrepreneurship in general, little is known about the *type* of skills that drive individual entrepreneurial performance (Hartog et al., 2010). Following the important contribution of Lazear (2005), (only) one skill-type related aspect has been studied rather rigorously: skill balance, or being a Jack-of-All-Trades (JAT). To be able to bring people, ideas and physical resources together, entrepreneurs must have knowledge, at least at a basic level, of a large number of business areas. This is required to successfully combine talents and manage those of others. Even though entrepreneurs can hire other people, the entrepreneur must possess a variety of skills to be able to judge other peoples' abilities and how to combine them. In this view, employees may be specialists, but entrepreneurs require a broad set of competencies in order to deal with the various tasks and challenges they are confronted with. Lazear's model predicts that entrepreneurs will benefit from having a balanced set of skills whereas employees will not.

Several empirical studies have tested Lazear's prediction.³ Most studies indeed find a positive relationship between individual balanced skills and entrepreneurial performance, where performance is measured either in terms of earnings (Hartog et al., 2010; Bublitz and Noseleit, 2014), new venture creation (Stuetzer et al., 2012, 2013; Hessels et al., 2014) and/or self-employment duration (Oberschachtsiek, 2012). One exception is the paper by Åstebro and Thompson (2011) who find that having a balanced skill set is driven by a taste for variety and thereby negatively related to earnings. Although, Lazear's theory has intuitive appeal (and empirical support)

¹See Klotz et al. (2014), Parker (2009), Lechler (2001) and Kamm et al. (1990) for an overview.

²More in general, there is convincing evidence that cognitive ability affects performance outcomes (Cutler and Lleras-Muney, 2010; Dohmen et al., 2010; Grinblatt et al., 2011), and determines — at the aggregate level — economic growth (Hanushek and Woessmann, 2008; Gennaioli et al., 2013).

³Lazear's Jack-of-all-trades theory has been tested in terms of entrepreneurship entry and performance. Positive results on entry have been found for example by Lazear (2005), Wagner (2006) and Lechmann and Schnabel (2011), while Silva (2007) shows that innate ability is more important than acquiring a balanced set of skills in the probability of becoming an entrepreneur. Given the focus of this paper we limit the discussion to the studies related to entrepreneurial performance.

when applied to solo entrepreneurs, its application to entrepreneurial teams is less clear cut. In this paper we test to what extent balanced skills are beneficial within entrepreneurial teams. Furthermore, we investigate whether a lack of balanced skills at the individual level can be substituted by combining the skills of different specialists within one team.

We conduct a field experiment to analyze the role and substitutability of balanced skills for teams with entrepreneurial tasks. We study teams of children who set up a toy business in friendship bracelets in an entrepreneurship education program (“BizWorld”) in the last grade of primary school in the Netherlands. The setting of the field experiment appears to be rather unusual at first sight but there are numerous advantages that will be presented below. Two basic skills are (most) important at this age: verbal ability and mathematical ability. Based on uniform and precisely measured scores on these skills for all 1,131 individuals, we exogenously compose 179 teams separated into four different team types: JAT-teams, (homogeneous) specialist teams and mixed specialist teams. JAT-teams consist only of JATs, i.e. children with balanced skills at the individual level. Additionally, within these teams the average scores on math and verbal ability are comparable. Math- and verbal-specialist teams consist of children who are either math- or verbal-specialists. These are children that score higher on one of the two subjects compared to the other. The mixed-specialist teams consist of a combination of math- and verbal-specialists. However, just as the JAT-teams, within these teams average math- and verbal ability (at the team level) are comparable. We compare the performance of the different team types to determine to what extent the measured skill sets affect team performance.

The results show that JAT-teams perform significantly better than both mixed-specialist teams and verbal-specialist teams. These results suggest that balanced skills are not only beneficial for individual entrepreneurs but also for entrepreneurs in teams. Despite the prediction of the JAT-theory, the performance of math-specialist teams is not significantly different compared to the JAT-teams. Moreover, our finding that mixed-specialist teams perform significantly worse than JAT-teams indicates that it is hard to substitute individual balanced skills by combining different specialists within one team. Apparently the ability to combine resources effectively is not something that comes across when people combine their specialized skills within teams.

Obviously, conducting a field experiment in the BizWorld entrepreneurship education program provides a trade-off between internal and external validity. However, this particular setting and the controlled experimental design have various benefits which allow us to study an interesting causal effect. First and foremost, there is no self-selection bias as all individuals are forced to start a business and do so in a team that is exogenously composed by us. When studying real life entrepreneurial teams the causal effect of team composition on team performance is difficult to determine, among others due to self-selection into and out of teams. Moreover, the possibility of running a field experiment in a setting where actual entrepreneurs are forced to start their business with a group of people randomly assigned to them is virtually impossible.

Furthermore, all children in our sample have approximately the same age and exactly the same level and amount of education. This is important for two reasons. First, because the accurate measurement of the sheer effect of cognitive skills is problematic when education levels and ages differ (Heckman et al., 2006). Age affects measured ability, whereas there is reverse causality in the relationship between education and measured ability, i.e. schooling affects test

scores and measured test scores predict schooling (Hansen et al., 2004, p.40). Second, the balancedness in skills for children at this age is exogenous, i.e. based on endowed skills instead of a selection of people who decided to invest in a diverse set of skills (Silva, 2007). Another benefit of this field experiment is that we are able to create a relevant and reliable measure of JATness. From the psychological literature we know that mathematical and verbal ability are the two most important cognitive skills for the learning outcomes of children aged 11-12 and in this experiment these skills are measured using a uniform and valid scale shortly before the children enter the program. Obviously, when studying venture teams in actual practice, each business would require different sets of skills and the set of relevant skills that must be balanced would not be so easy to define or measure in a uniform way.

Other advantages of using this program for the purpose of our field experiment are that the toy businesses are set up under identical circumstances and that we obtain uniform measures of performance that all become available at the same time. Thus, our data do not suffer from a 'survival of the fittest' bias. Finally, in this education program, the relationship and interaction between team members is a crucial component of team performance. All in all, the education program provides us with a suitable real effort experiment.

Our paper contributes to the existing literature in several ways. As discussed above, several studies have analyzed the implications of Lazear's JAT-theory for individual entrepreneurs. To the best of our knowledge, however, there is little evidence on the effect of balanced skills at the team level on (team) performance. Another contribution of our paper is that we explicitly study the intra-team substitutability of useful combinations of skills. Moreover, the use of a field experiment with randomized assignment to teams enables us to establish a causal effect of team composition on team performance.

The remainder of this paper is structured as follows. Theory and hypotheses are discussed in Section 2. Section 3 describes the entrepreneurship education program and its context. The research design is described in Section 4. Section 5 reports the empirical findings and in Section 6 we summarize and conclude.

2 Theory and hypotheses

The empirical studies discussed in the introduction examine the benefits of having a balanced skill set for individual entrepreneurs. Most studies find a positive link between balanced skills and individual entrepreneurial performance. However, these studies offer little insights into how balanced skills (within a team) could affect team performance. One of the predictions made by Lazear (2005) is that, as production processes become more complex, the supply of suitable entrepreneurs, i.e. individuals who have a balanced set of all the relevant skills, will decrease. Teaming up with others in order to overcome a lack of skill balance at the individual level might be one way to overcome this problem. Indeed, the number of team start-ups is increasing rapidly (Klotz et al., 2014). Therefore, it is important to gain insight in the effective composition of (entrepreneurial) teams.

Studying effective team composition in general is certainly not novel. The effectiveness of multiple aspects of team composition have been studied in various environments, albeit most

often in ways that do not permit causal inferences (Adams et al., 2010). There are notable exceptions. Most relevant for our study are papers that look at team composition in terms of skills. Kahane et al. (2013), Franck and Nüesch (2010) and Gould and Winter (2009) have studied the effect of skill dispersion on team outcomes using various professional sports settings (where the market is international and performance and skills are easily measurable). Moreover, Leonard et al. (2004) and Hamilton et al. (2003) have studied the effective composition in terms of skills relevant to production processes in retail and a garment factory, respectively.⁴

Entrepreneurship obviously requires a broad array of skills and involves complex problem solving and decision making. Several studies have been conducted to investigate the relationship between the composition of new venture teams and team performance (see Klotz et al. (2014) for an overview). Most of these papers study team diversity in terms of demographic characteristics (Chowdhury, 2005), functional background (Hmieleski and Ensley, 2007) and/or work experience (Eisenhardt and Schoonhoven, 1990). There are only a few studies that are related to the subject of interest in this paper, i.e. balanced skills within new venture teams. Colombo and Grilli (2005) show that there are possible gains to be achieved from the combination of economic-managerial and scientific-technical education among members within a start-up team. Cantner et al. (2010) study two different types of functional heterogeneity in new venture teams: knowledge scope and knowledge disparity. The results show that a broad knowledge stock is important for new venture growth but has no influence on survival. Similarity or functional overlap is positively associated with setting up and maintaining a new venture but has no influence on venture growth. Finally, Stuetzer et al. (2012) look at the relationship between balanced skills (measured by work experience in five different functional areas) and the progress of start-up projects by nascent entrepreneurs, both at the individual as well as at the team level. They find that balanced skills (for individuals and teams) are associated with making more progress in the venture creation process.

However, in all these studies the composition of the teams is endogenous, i.e. chosen by the members of the team. To the best of our knowledge, the only other paper that permits a causal inference of the effect of skills diversity on team performance is the study by Hoogendoorn et al. (2012). Their paper differs from this paper in terms of skill dimension. Hoogendoorn et al. (2012) look at diversity in one dimension, i.e. ability, and estimate the effect of dispersed levels of cognitive ability (or IQ) on team performance. The composition of the teams in our sample and the definition of skill diversity are based on two dimension, i.e. mathematical and verbal ability. For the team composition we take the levels of each of these skills into account as well as the difference between the two skills, both at individual and at the team level.

If balanced skills are needed in order to effectively deal with the various tasks and challenges that arise when setting up a business then, taken literally, Lazear's theory predicts that entrepreneurial teams consisting of a combination of individual JATs show better performance than teams consisting of individual specialists do. Because teams that have team members with

⁴Other examples of studies allowing causal inference of the effect on performance of team composition in other dimensions than skills are Adams and Ferreira (2009); Ahern and Dittmar (2012); Apesteguia et al. (2012); Dufwenberg and Muren (2006); Hoogendoorn et al. (2013) for gender. In addition the effect of team composition in terms of ethnicity, often in combination with some other demographic characteristics such as age and gender have been studied by Hamilton et al. (2003); Hansen et al. (2006) as well as by Kahane et al. (2013) and Hoogendoorn and Van Praag (2012).

only one type of specialized knowledge by definition lack some relevant skills needed for successful entrepreneurship. The theoretical model by Lazear (2005) for individual entrepreneurs, the empirical support from studies of individual entrepreneurship and some recent descriptive insights at the team level lead to our first hypothesis.

Hypothesis 1. *Entrepreneurial teams consisting of a combination of individual JATs perform better than teams consisting of (one type of) individual specialists.*

Yet, as an alternative to a team consisting of individual JATs, one could possibly combine the skills of different specialists within a team. It is possible that the balanced set of skills needed to be successful as an entrepreneur do not have to be endowed within one person, but can be spread out over several members of an entrepreneurial team. In this setting not all team members need to have all the skills as long as all the required skills are present within the team.

The literature reports both advantages and disadvantages of team diversity for team performance (Hamilton et al., 2003; Hoogendoorn et al., 2013; Kilduff et al., 2000). Possible gains from diversity could occur through complementarities and mutual learning (Lazear, 1999). However, these benefits might be offset by costs associated with communication and coordination which becomes more cumbersome in a diverse team (Lazear, 1999; Mäs et al., 2013; Richard et al., 2004). Several studies that looked specifically at start-up teams have shown the (net) benefits of diverse start-up teams. The presence of different types of skills within a founding team can produce multiple perspectives of a problem and thus can lead to more efficient solutions. Ensley et al. (2002) show that cognitive conflicts related to processes and ideas enhance entrepreneurial performance, as measured by new venture growth. Moreover, homogenous skills within a team can limit the creative process and thereby produce less innovative ideas (West, 2007). Hence it appears that there are certain benefits of diverse skills for successful entrepreneurship. However, the costs associated with team diversity have also been studied. In the management literature the division of teams into subgroups based on task related characteristics are defined as *informational faultlines* (Bezrukova et al., 2009). In a dynamic environment, such as the entrepreneurial process, strong informational faultlines may hinder quick and effective communication and coordination (Cooper et al., 2013). Moreover, if the differences between subgroups are (too) large, this could lead to tensions which could cause conflicts and thereby impede optimal performance (Mäs et al., 2013).

Thus, the mechanisms underlying possible performance differences between (the more homogeneous) JAT-teams and mixed specialist teams are twofold. First, JAT teams may have performance advantages over mixed specialist teams dependent on the substitutability of JATs by mixed specialists. This substitutability is unknown because it has not yet been studied. Second, the difference between JAT teams and teams of mixed specialists is a difference between homogeneity and diversity, which may be beneficial or costly. The net effect of diversity depends on how much the mixed specialists are able to benefit from the complementary skills of the other team members. In order for the benefits to be larger than the costs, they should be able to communicate effectively and learn from each other at low cost (Lazear, 1999). Taking both the costs and benefits of skill diversity into account, as well as the possible substitutability

of JATs by mixed specialists, it is not clear which effect will dominate in the comparison between JAT-teams and mixed-specialist teams. Therefore, we formulate our second hypothesis in terms of two alternative hypotheses.

Hypothesis 2a. *Entrepreneurial teams consisting of a combination of different types of specialists, that have comparable scores at the team level in the same relevant skills set, show lower performance compared to teams consisting of individual JATs.*

Hypothesis 2b. *Entrepreneurial teams consisting of a combination of different types of specialists, that have comparable scores at the team level in the same relevant skills set, show higher performance compared to teams consisting of individual JATs.*

3 Program and context

For our field experiment we use teams of children participating in one of the leading entrepreneurship education programs in the world, called the BizWorld program.⁵ The BizWorld Foundation was founded in the United States in 1997. Since then, over 450.000 children from 80 different countries have participated in one of their education programs.⁶ In the Netherlands the program started in 2004 and until now approximately 30.000 children have participated. We conducted our experiment in the spring of the school year 2009-2010 and included participating schools in our sample in the proximity of (our university in) Amsterdam in order to be able to monitor the schools closely.

The Dutch version of the program is taught by an entrepreneur (or someone from the business world) in cooperation with the class teacher. Prior to the start of the education program the teacher and the entrepreneur are introduced to each other during a train-the-trainer session. During this session the content of the program is explained and the teacher and the entrepreneur receive the "BizWorld-suitcase", containing a detailed course handbook and the course material (see Figure 1). The course material also included a letter for the parents informing them about the entrepreneurship education program and the related research project that was to be conducted. A researcher was present during each of the training sessions to clarify any questions that the teachers (or entrepreneurs) might have about the research project.

All Dutch primary schools -whether private or public- are eligible for BizWorld. Schools usually get in touch with the program through BizWorld marketing campaigns (i.e., BizWorld sending letters to schools to invite them to participate) or through sponsoring entrepreneurs or companies (from the neighborhood for instance). In general the BizWorld Foundation matches schools and sponsoring entrepreneurs willing to participate. Thus, financial or network constraints do not hinder schools' participation in the program.⁷

⁵The program and context described in this section are similar to those described by Rosendahl Huber et al. (2012).

⁶The BizWorld Foundation offers three different education programs for children from the third to the eighth grade (i.e. BizWorld, BizWiz and BizMovie). BizWorld is the largest with approximately two-third of all the children participating in this program (The BizWorld Foundation Annual Report, 2011).

⁷Concerning the representativeness of the sample (for the Dutch population of school kids in the last grade of

Figure 1: Course material



Schools sign up for the program at the beginning of each school year (before January). The children participating in the program are aged 11 or 12 and are in the last grade of primary school. Most schools have either one or two (parallel) classes in last grade. In general, the decision to participate is taken at the school level. The minimum level of participation is an entire class, i.e., individual pupils or teams cannot participate. The five-day program usually takes place within a period of 2 to 4 weeks during the last few months before the pupils go to high school. It is a structured program and a day-by-day overview of the content is shown in Table 1.

The first day provides a theoretical introduction to entrepreneurship and pupils learn about the basic concepts of running a business. Then, the class is divided into teams of 5 to 6 pupils (where each team serves as one company).⁸ Within each team each child can apply for a specific role. The positions that need to be filled are: CEO, CFO, Sales Director, Director of Product Design, Marketing Director and Director of Manufacturing. Each role comes with certain specific responsibilities, however, team work is crucial during the entire program.

On the second day, the teams decide on a company name and officially register their company with the "chamber of commerce".⁹ Next, all the teams write a business plan, which is presented to a "venture capitalist" in order to sell stocks and to raise start-up capital. The quality of the business plan together with the presentation determine the share price the investor is willing to pay. Each team receives ten shares at the registration of their company. They can decide on the number of shares they want to sell to the investor. However, each team has to sell some shares, because they need some cash before they can start the design and production process. During the program all transactions are made in fake, so-called BizEuros instead of actual Euros. On this day, the teams also prepare the design and a prototype of their product. The raw materials the teams can purchase are most suitable for the production of friendship bracelets, however other products such as key cords or bookmarks are also possible (see Figure 1).

primary school) we compare our sample to the population in terms of standard individual background characteristics, school characteristics and neighborhood characteristics. We collected information on those characteristics by means of the questionnaires, schools' websites and Statistics Netherlands, respectively (see Rosendahl Huber et al. (2012) for a more detailed description on the representativeness of the sample).

⁸In the regular course of the program the team composition is determined by the teacher. However, during our field experiment we were allowed to compose the teams. The details will be discussed in the next section.

⁹During the course of the program the children have to go to several official agencies, i.e., chamber of commerce, bank, venture capitalist, etc. These roles are played by the teacher.

Table 1: The BizWorld program

<i>Day 1</i>	Introduction and theory on entrepreneurship Apply for position in team
<i>Day 2</i>	Register company and receive 10 shares Present business plan to "venture capitalist" to raise start-up capital Company stock prices displayed in class
<i>Day 3</i>	Design and manufacture products (friendship bracelets) Calculate production costs (incl. rent, material, salaries, etc.) Determine product prices
<i>Day 4</i>	Design marketing campaign (poster and "commercial") Sell products to pupils in lower grade
<i>Day 5</i>	Complete profit- and loss statement and balance sheet Winning team announced and rewarded

It is important that everything is well planned and prepared for the production process which is scheduled for the third day. During the production phase the teams have a limited amount of time (one hour) to manufacture as many products as possible. When the production process is over, the total production costs are calculated. These include salaries of the directors, raw materials, rent, etc. On the basis of these costs, the teams determine the sales price of their products.

From the third day onwards, the teams have two alternative routes to raise more capital. They can either sell more shares to the venture capitalist, thereby reducing their ownership share in the company, or they can take up a loan from the bank which has to be redeemed, including interest, before the end of the program. It is explained to them that if they sell too many shares they lose ownership of the company. This is a difficult concept and the teacher and the entrepreneur try to convey as clearly as possible the importance of this factor in determining the winning team.

The fourth day is very dynamic. Teams design their marketing campaign, which consists of a poster, a slogan and a "commercial". The commercial can for instance be a two minute stage play or a recorded little commercial, dependent on the resources of the school and the ideas of the entrepreneur and teacher. Each team is given the opportunity to present their "commercial" in front of the group of prospective buyers before the "big sale" starts. The big sale is an organized fair at which the products are sold to the children in the grade below. After the sales market is over, the revenues are calculated.

On the last day of the program, each team has to complete a financial report consisting of a profit and loss statement and a balance sheet after having redeemed their loans and paid profit taxes. The financial report is a basic ingredient to assessing the performance of teams and to determine their *Team Company Value*. Based on this objective performance measure, teams are ranked within their class. At the end of this day the winning team is announced and rewarded.

Individual team members have strong incentives to care about the business performance of their team (Bradler et al., 2013). Firstly, the BizWorld foundation provides certificates for each member of the winning team. Furthermore, the entrepreneur or the company that sponsors the education program at the school offers some small prizes (usually in the form of gadgets) to

the winning team. Additionally, but only in the year we conducted the field experiment, team members of the winning team were awarded a gift voucher of € 7,50 each, and the team members of the runner up were each rewarded a gift voucher of € 5,00. The teacher could choose between two types of gift certificates; one that could be used at all major chain stores in the Netherlands or one that could only be spent on books. The majority of the teachers ($\frac{2}{3}$) chose the general gift certificate, the other teachers chose the book voucher. Finally, based on our observations, BizWorld teams usually show a strong motivation to achieve good company results and win the competition in their class.

4 Data and methodology

4.1 Design of the field experiment

The objective of this field experiment is to assess the effect of balanced skills for entrepreneurial teams. The use of the experimental setting in which teams of children perform entrepreneurial tasks within the setting of an entrepreneurship education program possibly imposes some limitations to the external validity of the results. However, it is not clear how findings on the effect of balanced skills on team performance for teams of children should differ from those found for adults. Moreover, the specific features of the environment and the controlled experimental design (which are described below in more detail), allow us to estimate an interesting and relevant causal effect.

The BizWorld program requires school children to set up a business in a team. The program also requires them to divide tasks by assigning distinct professional positions to each of the team members. Thus, this program provides us with an environment in which genuine team work is relevant and the task is entrepreneurship related

According to the psychological literature, verbal and mathematical skills are the two core skills for children between the age of 6 and 15 years.¹⁰ Arguably, all other skills, developed later in life, are derived from these two core skills. For instance analytical and technical ability are closely linked to mathematical ability, whereas communication and language skills can be associated with general verbal ability.

The measurement of these two skill types forms the core of a countrywide uniform exam that the children take just before the program starts. Hence, for both skills objective measures are available from this standardized test. The scores from these tests provide us with objective and comparable measures for (almost) all the children in the sample. At the stage of primary school where children have not yet been selected into various school types and levels, the variation in the scores of these skills is still maximal. Hence, based on an objectively measured and well defined skill set we are able to create a relevant and reliable measure of skill balance.

Moreover, the BizWorld program sets a clear and measurable objective for the participating teams, i.e., generating as much Team Company Value as possible and acquiring the highest ranking. An exact measure of this outcome variable is determined for all teams at the same

¹⁰In the commonly used Stanford-Binet test of abstract intelligence, verbal and quantitative reasoning (i.e. language and numerical ability) are two specific factors determining intelligence. Moreover, PISA (the OECD Program for International Student Assessment) distinguishes reading and mathematics as the two key subjects used for the worldwide evaluation of education systems (Source: www.oecd.org/pisa).

Figure 2: Team composition: from individuals to teams

Combination of skills (per person)			Individual type	in Team type
Math	Verbal			
H	H	= JAT _H	JAT	JAT-team
M	M	= JAT _M		
L	L	= JAT _L		
H	M	= Math _{HM}	math specialist	math-specialist-team
H	L	= Math _{HL}		
M	L	= Math _{ML}		
M	H	= Verb _{MH}	verbal specialist	verbal-specialist-team
L	H	= Verb _{LH}		
L	M	= Verb _{LM}		

*Note: For illustrative purposes we divided each skill into three possible levels: high (H), medium (M), low (L). In reality this is a continuum (0-100) of possible values.

stage (after having completed the program). Finally, for the sake of the experiment we were allowed to compose teams based on the scores of the two skill types.

We compose four distinct team types using a two-step procedure. First, the children are classified as being one of three individual types based on their abilities. A child is classified as either a JAT, a math specialist or a verbal specialist (see Figure 2). We define JATs as those children with equal mathematical and verbal ability. We thereby closely follow Lazear’s definition that JATs do not perform significantly better in one of the two tasks but they are equally good (or bad) in both (Lazear, 2005, p.656). Specialists are defined as pupils with a higher score in either math or verbal relative to the other subject.

For illustrative purposes, let us assume that for each skill there are only three possible levels: high (H), medium (M) or low (L), see Figure 2.¹¹ This would yield three types of JATs: high, medium and low ability and, similarly, three types of math- and verbal specialists. In a second step we use the individual types to compose the four different team types: JAT-teams, mixed-specialist-teams, math-specialist-teams and verbal-specialist-teams. As can be seen in Figure 2, the JAT-teams consist only of JATs and therefore have balanced skills both at the individual and at the team level, i.e. within these teams the average scores on math- and verbal ability are comparable. The mixed-specialist-teams, combining math- and verbal specialists, have comparable scores in math and verbal ability at the team level, but not at the individual level. Math-specialist-teams are composed of only math specialists and verbal-specialist-teams consist only of verbal specialists. Hence, these teams have a relatively high math or verbal score, both at the individual as well as at the team level average. We further make sure that the teams within one class are comparable in terms of average ability and gender composition. Otherwise the assignment of children to teams is random.¹²

4.2 Sample

The main source to measure math and verbal ability is a nationwide exam called the "CITO"-test. This exam consists of two mandatory parts, measuring math and verbal ability, respec-

¹¹In reality we measure a continuum (0-100) of possible values. The measurement of the skills in our sample is described in more detail in Section 4.2.

¹²Per class we tried to form as many "usable" teams as possible, i.e., teams that fit the team composition shown in Figure 2. However, it also occurred that there were some children left that couldn’t form a team fitting the rules. We placed these children together in one team and labeled this as "leftover"-teams.

Table 2: Overview of test score types

Test score type	Pupils		Teams	
	#	%	#	%
CITO (<i>standardized test score</i>)	709	48	124	48
CITO (<i>% of correct answers</i>)	208	14	36	14
LVS	248	17	43	17
DO	110	7	18	7
School grades	57	4	10	4
Unavailable	155	10	25	10
Total	1487	100	256	100

tively.¹³ We received these CITO math and verbal scores from the schools for the majority of the pupils in our sample (i.e., 62%), see Table 2. The scores were reported by the schools in two different formats: the official, standardized scores (48%) ranging from 0 to 100, where 50 corresponds to the nation wide average or, alternatively, the percentage of correct answers on both the math and verbal test per child (14%).

Not all schools in our sample participate in the CITO-test.¹⁴ Almost all the schools that did not participate in this exam provided us either with the grades from a standardized student tracking system, called the "Leerlingvolgsysteem" (LVS) or with the scores from another type of nationwide exam, called the "Drempelonderzoek" (DO). The LVS records the pupil's progress from the first until the last grade of primary school. For each subject several standardized tests are conducted during each school year where test scores range from A to E. We use the math and verbal test scores from the LVS for 17% of the pupils in the sample. We obtained scores from the DO-test for 7% of the pupils in the sample. The DO-test is comparable to the LVS and CITO-test, with test scores ranging from 65 to 135, where 100 corresponds to the nationwide average. The scores from both the CITO and the DO-test can be converted directly into grades that correspond to the grades from the LVS. Finally, 10% of the schools (and pupils) did not provide any test scores (on time) and were removed from the sample. An overview of the number of pupils and teams per type of test score is shown in Table 2.¹⁵

Besides collecting test scores from the schools, we obtained information directly from pupils by means of two extensive questionnaires, one prior to the start of the program (pre-treatment) and one after the program (post-treatment). The overall response rate for both questionnaires is 92,5%.¹⁶ The pre-treatment questionnaire contains a wide variety of questions on individual background characteristics, such as age, gender, ethnicity, occupational status of the parents, etc. Additionally, both questionnaires contain some questions to assess the children's knowledge on entrepreneurship before and after the program. At the team level the post-treatment questionnaire is used to collect some information on team characteristics, such as the number

¹³Besides, schools have the option to include a part testing 'world orientation', which is a combination of history, geography and biology.

¹⁴In the Netherlands approximately 80% of the primary schools participates in the math and verbal test of this exam.

¹⁵One school did not provide us with test scores from any of the above mentioned standardized tests, but gave us the students' grades (A-F) from a school exam instead.

¹⁶One class did not return the second questionnaire and for some children one of the questionnaires is missing due to absenteeism on one of the test days.

of conflicts within a team and (self-assessed) team work. We have developed and tested these questionnaires in close collaboration with a child psychologist.

Data on team performance are obtained via the teachers. They filled out a standardized spreadsheet during and at the end of the program to register all transactions made by the teams, such as number of shares sold, share price, revenues etc. We use this spreadsheet to make sure that all the teachers collect the same information and calculate the team performance in exactly the same way.¹⁷ The teams also registered all transactions themselves and completed a financial overview of their company at the end of each day. We consider the information collected by the teacher as providing the more objective and accurate results at the end of the program (teams turned out to sometimes make mistakes or miscalculations). This information is used to determine the winning team and the ranking of teams. The response rate for the completed excel files was 87,3% (i.e. for 8 out of 63 classes we did not receive the excel file). The final sample for which we have received all the required information, i.e. both the math- and verbal scores and the results on team performance, consists of 1131 pupils in 179 teams.¹⁸

4.2.1 Individual types

A few weeks prior to the start of the education program we received the names of the children, their gender and their test scores from one of the tests described in Section 4.2. We measure the balance in mathematical and verbal ability by taking the (absolute) difference in the test scores between the two subjects. To define an individual as a JAT, we use a maximum difference of 15 (percentage) points as a cutoff point for the test scores from the nation wide exams (i.e. CITO- and DO-test). This cut-off point is chosen because it is the smallest unit of distinction for the grades in the LVS, i.e., a 15 point difference in test score corresponds to a one grade difference in the A-E scores of the LVS (and the grades from the school exam).¹⁹ Specialists are defined as those children with a difference between the two test scores of more than 15 (percentage) points in the nationwide exam or a difference in grades of at least 1 for the LVS or the school exam. The use of this definition implies that the group of specialists also includes some pupils that are not very good in either of the two subjects. However, these specialists still have a comparative advantage in one subject. Based on this choice of cut-off points, our sample consists of 720 JATs, 292 math-specialists and 306 verbal-specialists. The left hand side of Figure 3 shows the distribution of the individual types in the sample.²⁰

4.2.2 Team types

The individual types are used as the basis for our team composition (see Figure 2). Because participation in the education program is at the class level, the assignment of individuals to teams takes place within classes. Per class, the possibilities depend on the number of children

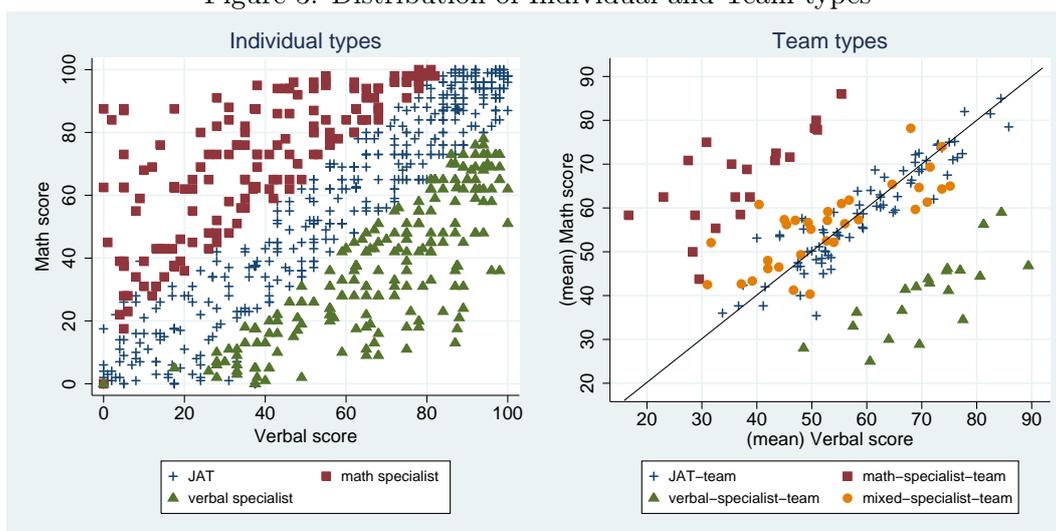
¹⁷An example of the teacher spreadsheet is shown in Figure A1 in the Appendix.

¹⁸We compared the initial and final sample in terms of their individual and team characteristics and confirmed that there is no indication of non-random attrition based on observable characteristics.

¹⁹For the DO scores, we used an official conversion table to match the scores to the grades of the LVS to confirm that the classification of the individuals based on the DO scores is accurate (see: <http://www.drempelonderzoek.nl>).

²⁰Besides the 155 pupils for whom the test scores for the entire class are missing, 14 individual test scores are missing of pupils spread out over different classes. Hence, these children are assigned to "leftover"-teams.

Figure 3: Distribution of Individual and Team types



(on average 24) and the distribution of individual types. Most teams consist of five or six pupils.²¹ Two considerations guide the otherwise random allocation of children to teams within each class given their individual type: teams should be equal as much as possible in terms of average team ability and gender composition.²²

After the categorization of the individual types in each class is done, the JATs are assigned to the JAT teams (see Figure 2) and then we form (basically) either mix or separate the specialists (this choice is randomly distributed over classes). The resulting sample of teams consists of 117 JAT-teams, 41 mixed-specialist-teams, 23 math-specialist-teams and 27 verbal-specialist-teams (see the right hand side of Figure 3). The other 23 teams were ‘leftover’ teams and consisted of combinations of individual types that couldn’t be classified as any of the team types of interest.

The teachers were not informed about the details and the purpose of the team composition. We merely informed them about the resulting team compositions and that the teams should not be changed without our prior consent, unless they had strong objections against the team assignment. Based on the teachers’ objections, 20 children moved teams prior to the start of the program. As a result, five teams were no longer usable. A researcher visited each school at the end of the education program to confirm that no changes had been made to the initial team composition of the usable teams. The final number of teams and the descriptives of some of the main characteristics per team type are shown in Panel A of Table 3.

To test for pre-treatment differences, we estimate whether team types differ in terms of their relevant average team background characteristics, see Table 3. As intended, given our experimental design, the teams differ in terms of average mathematical and verbal ability: math-specialist teams have a significantly higher math score and verbal-specialist teams have a significantly higher verbal score compared to the benchmark of JAT teams. Moreover, and in line with nation wide averages, girls in our sample score higher on verbal ability and boys score higher on

²¹There are 18 teams with seven team members and one team of only four children.

²²A pilot study we conducted in the year prior to this experiment revealed that girls have a comparative advantage in the production process (of friendship bracelets, key cords, etc.). Ability is defined here as the unweighted average of the sum of the two subjects. Otherwise none of the individual non-ability related characteristics was associated with team outcomes.

mathematical ability. Therefore, despite our efforts to create balanced teams in terms of gender composition, the average share of females is significantly higher in verbal-specialist teams and significantly lower in math-specialist teams. In the same vein, even though our aim was to create teams of similar average ability, the (average) ability levels are significantly lower for mixed-specialist and verbal-specialist teams compared to JAT- and math-specialist teams. We will control for these differences in our final estimations.

4.3 Outcome variables

For the evaluation of the effect of team composition on team performance, we use the following outcome measures. The first (and main) outcome measure is the team’s ranking within their class. For the majority of the classes ($\frac{2}{3}$) the ranking of the team is based on the financial measure *Value of own shares*. This is calculated as total company value multiplied by the fraction of shares still owned by the team (i.e. not sold to the investor). For the remaining 33% of the teams in the sample the ranking equals the (unweighted) average of the ranking based on the *Value of own shares* and the ranking based on the number of *Sustainability Points*. The fact that there is no uniform performance measure underlying the team ranking is not problematic because the measure is uniform within each competitive environment, i.e. within one class.

On average there are 4.62 teams per class, with a minimum of two and a maximum of nine teams per class.²³ Hence, team ranking can vary between one and nine, 1 being the best and 9 being the worst. Because it is easier to win in a class with a few teams than in a class with many teams, we divide the rank of each team by the number of teams in its class such that Team Rank is normalized between 0 and 1. We consider this *Normalized Team Rank* as our main outcome variable.

Additionally, we will use *Value of own shares* as an outcome measure for the sub sample of teams where this was the only measure to determine team ranking. This performance measure provides more detailed information on team performance than only ranking. The sub sample consists of 113 teams (61 JAT-teams, 25 mixed-specialist-teams, 14 math-specialist-teams and 13 verbal-specialist-teams) and is thus large enough to make useful estimations.

As a final, alternative measure of team performance we will use the *Money won in the Tournament*. The ratio of the prize money that can be won by the members of the winning team and the members of the team that comes in second place is 3:2. There was no money to be won by any of the other teams in the class. Hence, we assigned the value three (3) to all the winning teams, the value two (2) to the second best teams and zero (0) to all the remaining teams. An advantage of this outcome measure is that it takes the incentives to become first or second in the tournament into account. However, it does not take into account the non-financial benefits of ranking highly, nor does it control for the size of the competition, i.e. the number of teams in the class. The descriptive statistics of the three outcome measures are shown in the lower panel (B) of Table 3.

²³There is just one class with these two extreme values. More in general, there are eight classes with less than four teams per class and three classes with more than six teams.

Table 3: Descriptives by team type

	Total sample		JAT-teams		mixed-specialist-teams		math-specialist-teams		verbal-specialist-teams	
	mean	std. dev.	mean	std. dev.	mean	std. dev.	mean	std. dev.	mean	std. dev.
A: Background characteristics										
Math score	56.86	12.84	59.05	11.49	56.53	9.16	67.42**	11.54	40.34***	9.22
Verbal score	57.13	13.82	59.67	11.30	53.09***	12.53	39.16***	11.18	69.48***	9.29
Δ (Math - Verbal)	10.41	11.58	3.28	3.00	6.55***	5.20	28.26***	6.94	29.14***	5.28
% Female	0.51	0.15	0.52	0.13	0.47*	0.12	0.38***	0.16	0.65***	0.18
Age	11.73	0.33	11.72	0.33	11.70	0.31	11.85	0.37	11.73	0.36
Average ability	3.07	0.71	3.20	0.69	2.88***	0.76	3.09	0.64	2.76***	0.65
Team size	5.76	0.61	5.77	0.64	5.89	0.45	5.47**	0.52	5.67	0.73
Nationality parents: One non-dutch	0.14	0.16	0.15	0.16	0.13	0.15	0.08**	0.12	0.18	0.19
B: Outcome variables										
Normalized Team Rank	0.57	0.27	0.53	0.27	0.64*	0.28	0.49	0.19	0.72***	0.22
Money won in Tournament	1.23	1.30	1.33	1.35	1.08	1.26	1.73	1.16	0.62***	1.02
Value of own shares	109.35	100.95	122.31	111.57	91.28	81.49	130.57	104.41	60.38***	55.85
Number of teams	179		105		38		15		21	

*Note: */**/** indicates a significant difference at the 10%/5%/1%-level in the team level average of the variable compared to the JAT-teams. Math- and verbal score and the absolute difference between the two scores are based on the nationwide exam (CITO) and the converted scores from the student tracking system (LVS). These are available for 150 teams. Average ability is proxied by the intended future high school track and ranges from 1 (pre-vocational secondary education) to 5 (pre-university education). Normalized Team Rank is the rank per team divided by the number of teams in its class. Money won in Tournament is the ratio of the prize that can be won by the winning team, the runner up and all the other teams (3:2:0). The values for Value of own shares are based on the sub-sample of teams where this was the measure to determine team ranking (N=113).

4.4 Randomization

There are several issues that could affect the design of our field experiment and the internal validity of our findings in a negative way. One of the main concerns is non-random attrition due to a non-random selection of teams or classes that failed to report their results. To test for this, we regress an indicator for whether or not the team results were missing on the different team type dummies and on various background characteristics at the team level. The results from these estimations show that attrition was indeed random. The internal validity of our results further hinges on the random assignment of individual types to the various team types. This issue is only relevant for specialists who can be assigned to either mixed or unmixed teams of specialists. JAT-individuals, on the contrary, can only be assigned to JAT-teams. We tested whether the assignment of specialists to team types was random by running regressions at the individual level for both types of specialists separately. The dependent variable is a dummy variable indicating whether an individual is assigned to a mixed-specialist-team (value 1) or to a specialist team matching his/her own type (value 0), whereas the independent variables are individual characteristics such as age or gender. The estimation results in Table A1 in the Appendix indicate that the assignment was indeed random.²⁴

Finally, we check whether the changes to the team composition that were made on the teacher's request after the initial assignment (but prior to the start of the program) affect our research design. We do this by comparing the (observed) team characteristics of the initial teams with the teams in our final sample. The results from this comparison show no systematic (significant) differences.

5 Results

5.1 Main results

To test our hypotheses and to estimate the effect of the different team types on team performance we regress each of the outcome variables on the team type dummies, using the JAT-teams (i.e. the largest group in our sample) as the benchmark. We control for several team and class level characteristics.²⁵ The team level characteristics we include are average age, gender composition (i.e. the share of females in the team), team size and average team ability. At the class level we control for the total number of teams per team type per class. The rationale behind these class level controls is that the team composition as well as the competition between teams took place at the class level. As a result, in some classes all the different team types are represented in the tournament, whereas in other classes certain team types are missing or might be overrepresented. Thus, by including the sum of teams of each team type per class, we account for the differences in the tournament composition across classes. The observations are clustered at the class level to obtain robust standard errors to account for the possibly correlated performance of the teams

²⁴Two differences are (only) marginally significant. Mixed-specialist teams have slightly fewer female math specialists and math-specialists have slightly higher math scores than those in pure math-specialist teams. Given our effort to compose balanced teams in terms of gender and ability, none of these two differences are surprising because verbal specialists are more likely to be female and of a slightly lower ability level (see Table 3).

²⁵The results for specifications without these controls are similar though, and can be found in Table A2 of the Appendix.

Table 4: The effect of team type on team performance

	Normalized Team Rank		Value of own shares		Money won in Tournament	
	(1)	(1)	(2)	(2)	(3)	(3)
math-specialist-team	0.01	(0.07)	-4.20	(27.90)	0.23	(0.37)
mixed-specialist-team	-0.12*	(0.07)	-50.48**	(23.98)	-0.29	(0.28)
verbal-specialist-team	-0.23***	(0.07)	-60.19***	(19.58)	-1.21***	(0.31)
<i>(omitted category: JAT-team)</i>						
Team characteristics						
Age	0.17***	(0.05)	41.28	(29.59)	0.59***	(0.22)
Female	0.03	(0.15)	-125.85*	(72.89)	0.53	(0.64)
Average ability	0.01	(0.02)	14.54	(18.33)	-0.04	(0.11)
Team size	0.06	(0.03)	24.22	(20.82)	0.15	(0.16)
Class characteristics						
<i># of teams per class:</i>						
JAT-teams	0.002	(0.01)	-20.79	(16.89)	-0.30***	(0.05)
math-specialist-teams	0.03	(0.03)	-12.50	(27.79)	-0.26*	(0.14)
verbal-specialist-teams	0.07***	(0.02)	23.18	(16.15)	0.15	(0.10)
mixed-specialist-teams	0.03	(0.02)	14.17	(17.00)	-0.20**	(0.08)
leftover-teams	0.02	(0.02)	-6.31	(21.54)	-0.16	(0.10)
Constant	-2.93***	(0.68)	-447.53	(474.49)	-5.47*	(2.91)
Number of teams	178		112		178	
R^2	0.14		0.17		0.17	

* Note: */**/** indicates a significance at the 10%/5%/1%-level. Observations are clustered at the class level. Robust standard errors in parentheses. Average ability is proxied by the intended future high school track and ranges from 1 (pre-vocational secondary education) to 5 (pre-university education). The results for Value of own shares were obtained using a sub sample for which this outcome measure was the only determinant for team ranking.

within one class.

The estimation results are shown in Table 4. To simplify the interpretation of the estimated coefficients for Normalized Team Rank, we multiply these values by -1. The estimated coefficients indicate the difference in performance of each team type compared to the JAT-teams (with robust standard errors in parentheses). Each column shows the result for one of the (three) outcome variables. When comparing the results in the three columns in Table 4, a consistent pattern emerges. JAT-teams and math-specialist teams perform equally well in terms of all three outcome variables, i.e. *Normalized Team Rank*, *Value of own shares* and *Money won in the Tournament*. This finding is partly in line with our first hypothesis, i.e., that JAT-teams would outperform teams consisting of one type of individual specialists. The hypothesis holds for the verbal-specialist teams but not for the math-specialist teams.

Furthermore, we find that the performance of the JAT-teams is significantly higher than the performance of the mixed specialists. These results indicate that the ability to combine resources effectively is not something that comes across when people combine their specialized skills within teams. Note that we obtain a negative coefficient for the mixed-specialist-teams for all three outcome variables. For two of the outcome variables, i.e. *Normalized Team Rank* and *Value of own shares*, this difference is statistically significant (p -values: 0.09 and 0.04). For the third outcome measure, *Money won in the Tournament*, the coefficient is not significantly different from zero. Thus, the performance of the mixed-specialist-teams is on average lower than the

performance of the JAT-teams. These results support hypothesis 2a. This indicates that it is not possible to compensate the lack of skill balance at the individual level by combining (the skills of) two types of specialists within one team.

Finally, the setup of the experiment allows us to analyze the effect of the two specific skills, i.e. math and verbal ability, separately. As stated above, math-specialist-teams perform equally well as JAT-teams. Thus, in line with the findings by Hartog et al. (2010) for solo entrepreneurs, we find that mathematical ability is also beneficial for entrepreneurial teams (some additional evidence is found in Van Praag and Cramer, 2001). The equal performance of JAT-teams and math-specialist teams suggests some substitutability between mathematical ability and balanced skills. Math-specialist teams have significantly higher mathematical ability than the JAT-teams (see Table 3). Yet, JAT teams are able to make up for this lack in mathematical ability by the presence of individuals with more balanced skills within their team. Alternatively, one could argue that math-specialist teams compensate the missing skill balance by sufficient mathematical ability. Nevertheless, it appears that team composition is important. Mixed-specialist teams have the same mathematical ability and the same skill balance at the team level as the JAT-teams, but show lower performance. Thus, the benefits of skill balance only exist if a team consists of only JATs. Moreover, if we control for average math ability and the highest math score within the team, the coefficients of these variables are insignificant and the results remain the same (albeit slightly less significant due to the loss of observations). This indicates that the results are not driven by just one high ability math person or superstar.

The effects of verbal ability are very consistent, and (relatively) negative, for all three outcome variables. Table 4 shows that the performance of the verbal-specialist-teams is significantly lower compared to the JAT-teams for all three outcome measures. This is also consistent with the findings by Hartog et al. (2010) that verbal ability does not yield a positive income effect for (individual) entrepreneurs.

The class or tournament characteristics provide some (additional) support for our findings. The estimations show that a larger number of JAT-teams or math-specialist-teams in the class is negatively related to team performance in terms of *Value of own shares* and *Money won in the Tournament*. Additionally, we find that a larger number of verbal-specialist-teams in the class is positively related to (all three) outcome measures (although only significant for *Normalized Team Rank*). This is exactly what one would expect based on the main effects: JAT teams as well as teams consisting of math specialists provide fiercer competition because they outperform the others.

5.2 Mechanisms

Given the results we find, it would be valuable to be able to address the issue of *why* teams with JATs and math specialists perform better than teams of mixed or verbal specialists.

As discussed in Section 2 there are both costs (e.g. communication and coordination costs) and benefits (e.g. increased creativity and mutual learning) associated with team diversity. We can measure coordination and communication costs in terms of the number of conflicts in a team and team work, where the latter is an inverse measure of these costs (similar to Hoogendoorn and Van Praag, 2012). These are based on two team variables measured in the post-treatment

questionnaire for this purpose: the number of conflicts within a team and team work. Both measures are self-assessed. The number of conflicts is measured in terms of the average of the number of reported conflicts by individual team members within one team. Team work is measured by the (team average) answer to the question: "How well did you work together?". We measure this on a 5-point scale where 1 = very well and 5 = very bad.

Comparing the average number of conflicts per team across team types, we see that there are indeed more conflicts in the mixed-specialist teams than in the (more homogeneous) JAT-teams. However, this difference is not statistically significant. Team work, our second measure, might also be easier accomplished in less diverse teams. The differences in team work across team types show that mixed-specialist teams (and verbal-specialist) teams score (marginally) significantly lower compared to JAT-teams (p -value: 0.08).²⁶ Hence, we find some evidence of higher coordination and communication costs in teams of mixed specialists compared to JATs.

These costs may potentially explain the better performance of JAT teams compared to mixed specialists if we also find a direct effect of these costs on performance. However, the design of our experiment does not allow any causal inferences other than related to team type and team performance due to a lack of other exogenous variation.²⁷ For instance, the self-reported measures of conflicts and team work might be affected by the performance of the teams, rather than the other way around, where the winning teams judge their team work more positively ex-post than the members of the losing teams. A factor that limits the applicability of this explanation is that the (homogeneous) teams of verbal specialists also have lower scores on team work than JAT and math-specialist teams. This is at odds with the explanation that diversity would cause higher coordination and communication costs.

Our final exercise to find suggestive evidence of what underlying mechanism might explain our results is related to the benefits side of more diversely composed teams. More diversely composed teams would have more complementary skills (which we imposed on the mixed-specialist teams but not on the others) and this might result in better (mutual) learning outcomes. To test this assertion we regress the development of entrepreneurship knowledge and nine non-cognitive skills typically associated with entrepreneurship on the team type dummies. We do not find any differences in skill development across the different team types.

All in all these results can only provide suggestive evidence about the association between team *diversity* and performance. The distribution over team types of the self-assessed numbers of conflicts within the team and the quality of the team work suggests that some costs of diversity are associated with mixed-specialist teams. However, we find the same disadvantage for the relatively homogeneous teams of verbal specialists. Therefore, we conclude that it is somewhat unlikely that the greater *diversity* of the teams of mixed specialists causes their lower performance relative to JATs.

²⁶Once we control for team characteristics such as team size, gender composition and average ability, the coefficients remain negative but they are no longer significant for mixed-specialist teams.

²⁷We find an insignificant relation between the average number of conflicts in a team and performance and a significant and positive relation between team work and performance (according to all three performance measures).

5.3 Robustness checks

We have performed several checks to confirm the robustness of our findings. Firstly, the results are robust to the in- or exclusion of team and class characteristics as control variables. Secondly, we ran separate regressions excluding the teams from our sample that were changed, upon the teacher’s request, after the initial assignment. Excluding the five affected teams from our estimation sample in which an individual team member was replaced by another, but where the team type remained unchanged, leaves the results qualitatively the same.

To make sure that the results we find are not driven by the four JAT-teams with very high math (and overall) ability (see Figure 3), we exclude those four teams from the sample. The results from these estimations are similar to the main results presented above, i.e. the performance of mixed-specialist teams compared to JAT-teams is lower in all three specification, but is only (marginally) significant for the outcome measure *Value of own shares*.

In another robustness check we test whether the results are not driven by a few superstars within the teams. To do this we create dummy variables of being a superstar for those individuals at the top 10% of the ability distribution for verbal and mathematical ability respectively. More specifically, since mathematical ability appears to be important at the team level, we test if performance is different for teams with a *math superstar* by including an interaction term of the *math superstar* dummy and the different team types. The coefficients for these interaction terms are insignificant in all specifications, so there appears to be no additional benefits from having a *math superstar* in the team.

An alternative explanation for the high performance of the JAT-teams compared to mixed-specialist teams is that, by definition, only JAT-teams include pupils who score very high on both math and verbal ability. *JAT superstars* are individuals whose test scores for both math and verbal ability are in the top 10% of the distribution (about half the JAT-teams has a *JAT superstar* according to this definition). If the performance of the JAT-teams is driven by only a few superstars, then our findings are not the results of team work, but rather the result of one high ability JAT person leading the team to high performance. To test if having a *JAT superstar* affects the performance of the JAT-teams, we use the subsample of JAT-teams and regress team performance on a dummy variable indicating if a team has one (or more) JAT superstars in the team, including the other controls we use in the main estimation. The coefficient for this dummy is not significant in any of the specifications, indicating that the performance of the JAT-teams is not driven by just one superstar in the team.

Finally, we estimate our main specification(s) without the teams that do not provide us with the (official) test scores from the nationwide exam or the standardized student tracking system. That is, we excluded those classes for which we received less precise measures, i.e., the percentage of correct answers from the nationwide exam or the grades of school exams (see Table 2). The results are very similar, although somewhat less significant due to the loss of observations.

6 Summary and conclusion

Team entrepreneurship is a growing phenomenon. In this study we focus on the skill composition of successful teams performing entrepreneurial tasks. Successful entrepreneurs are vital for

economic outcomes and empirical research has shown that the performance of entrepreneurs is determined by skill rather than luck. However, little is known about the *type* of skills that drive (individual) entrepreneurial performance in general and about the skill composition of successful entrepreneurial *teams*.

For solo entrepreneurs one skill-type related aspect has been studied more systematically: skill balance, or being a Jack-of-All-Trades (JAT) appears beneficial for entrepreneurs, in line with Lazear's JAT theory (2005). In this paper we empirically explore the extent to which skill balance is also important for entrepreneurial teams. To this end we test two hypotheses. The first is whether teams consisting of a combination of individual JATs show better performance than teams consisting of individual specialists do. The second hypothesis investigates whether it is possible to substitute (a lack of) individual balanced skills by combining the skills of different specialists within one team.

To test these hypotheses we have conducted a field experiment. We find an environment to analyze the role and substitutability of balanced skills for teams with entrepreneurial tasks within the rather unusual setting of an entrepreneurship education program ("BizWorld") in primary schools in the Netherlands. Teams of 5-6 children in the last grade set up a toy business in friendship bracelets. Two skills are relevant at this age: verbal and mathematical ability. Based on uniform and valid measures of the scores on these skills, we compose four different team types: JAT-teams, verbal-specialist teams, math-specialist teams and mixed-specialist teams. These teams set up a business in a uniform and competitive environment and we obtain uniform measures of the teams' performance.

Comparing the performance of the different team types we find that JAT teams do (weakly) better than teams consisting of individual specialists. In particular, JAT-teams perform significantly better than both mixed-specialist teams and verbal-specialist teams. Math-specialist teams perform equally well. Taken together these results partly support our first hypothesis, indicating that balanced skills are not only beneficial for individual entrepreneurs but also for entrepreneurs in teams. Moreover, based on these findings we find support for our alternative hypothesis 2a. The lower performance of the mixed specialist teams indicates that it is hard to substitute individual balanced skills by combining different specialists within one team. However, more research is needed to determine the exact mechanism through which this lower performance occurs.

Obviously, this experimental design has some limitations. The most important limitation is the possibly limited external validity of our results. The fact that we study teams of children participating in an entrepreneurship education program instead of actual entrepreneurial teams, poses limitations to the generalizability of our results. However, (a priori) there is no clear reason to assume that the treatment effects found for children should vary substantially with age or subject pool. Another possible limitation is that we define JATs and specialists only in terms of two types of skills: mathematical and verbal ability. As we noted already these are considered the two core skills for children at the age of 12 and form the basis of relevant skills to be developed later in life, according to the psychological literature.

Besides these limitations there are some notable advantages to the use and current set up of our field experiment which allows us to measure the performance effect of balanced skills with

a high level of internal validity. Most importantly, there is no self-selection bias, neither into or out of entrepreneurship in general, nor into or out of doing so in a team. So far, as discussed in Section 2, few empirical studies on effective team composition have taken into account that team composition is non-random and that the performance effects that are measured might be biased as a consequence. The studies that do allow causal inferences have not taken place in environments requiring a broad array of skills and involving complex problem solving and decision making, which is a realistic feature of influential teams, such as management boards or teams of entrepreneurs. The entrepreneurship education program provides us with a suitable real effort team task that arguably has these characteristics. Thus, this study provides an important first step in unraveling the effective composition of entrepreneurial teams.

References

- Adams, R. and Ferreira, D. (2009). Women in the boardroom and their impact on governance and performance. *Journal of Financial Economics*, 94(2):291–309.
- Adams, R., Hermalin, B., and Weisbach, M. (2010). The role of boards of directors in corporate governance: A conceptual framework and survey. *Journal of Economic Literature*, 48(1):58–107.
- Ahern, K. and Dittmar, A. (2012). The changing of the boards: The impact on firm valuation of mandated female board representation. *The Quarterly Journal of Economics*, 127(1):137–197.
- Apestequia, J., Azmat, G., and Iriberry, N. (2012). The impact of gender composition on team performance and decision making: Evidence from the field. *Management Science*, 58(1):78–93.
- Åstebro, T. and Thompson, P. (2011). Entrepreneurs, jacks of all trades or hobos? *Research Policy*, 40(5):637–649.
- Bezrukova, K., Jehn, K., Zanutto, E., and Thatcher, S. (2009). Do workgroup faultlines help or hurt? a moderated model of faultlines, team identification, and group performance. *Organization Science*, 20(1):35–50.
- Bradler, C., Dur, R., Neckermann, S., and Non, A. (2013). Employee recognition and performance: A field experiment. *Tinbergen Institute Discussion paper*.
- Bublitz, E. and Noseleit, F. (2014). The skill balancing act: when does broad expertise pay off? *Small Business Economics*, 42(1):17–32.
- Cantner, U., Goethner, M., and Stuetzer, M. (2010). Disentangling the effects of new venture team functional heterogeneity on new venture performance. *Jena Economic Research Papers*, 029.
- Chowdhury, S. (2005). Demographic diversity for building an effective entrepreneurial team: is it important? *Journal of Business Venturing*, 20(6):727–746.
- Colombo, M. and Grilli, L. (2005). Founders human capital and the growth of new technology-based firms: A competence-based view. *Research Policy*, 34(6):795–816.
- Cooper, D., Patel, P., and Thatcher, S. (2013). It depends: Environmental context and the effects of faultlines on top management team performance. *Organization Science*, forthcoming:1–20.
- Cutler, D. and Lleras-Muney, A. (2010). Understanding differences in health behaviors by education. *Journal of Health Economics*, 29(1):1–28.
- Dohmen, T., Falk, A., Huffman, D., and Sunde, U. (2010). Are risk aversion and impatience related to cognitive ability? *The American Economic Review*, 100(3):1238–1260.
- Dufwenberg, M. and Muren, A. (2006). Generosity, anonymity, gender. *Journal of Economic Behavior & Organization*, 61(1):42–49.

- Eisenhardt, K. and Schoonhoven, C. (1990). Organizational growth: Linking founding team, strategy, environment, and growth among us semiconductor ventures, 1978-1988. *Administrative Science Quarterly*, 35(3):504–529.
- Ensley, M., Pearson, A., and Amason, A. (2002). Understanding the dynamics of new venture top management teams: cohesion, conflict, and new venture performance. *Journal of Business Venturing*, 17(4):365–386.
- Franck, E. and Nüesch, S. (2010). The effect of talent disparity on team productivity in soccer. *Journal of Economic Psychology*, 31(2):218–229.
- Gennaioli, N., La Porta, R., Lopez-de Silanes, F., and Shleifer, A. (2013). Human capital and regional development. *The Quarterly Journal of Economics*, 128(1):105–164.
- Gompers, P., Kovner, A., Lerner, J., and Scharfstein, D. (2010). Performance persistence in entrepreneurship. *Journal of Financial Economics*, 96(1):18–32.
- Gould, E. and Winter, E. (2009). Interactions between workers and the technology of production: Evidence from professional baseball. *The Review of Economics and Statistics*, 91(1):188–200.
- Grinblatt, M., Keloharju, M., and Linnainmaa, J. (2011). Iq and stock market participation. *The Journal of Finance*, 66(6):2121–2164.
- Hamilton, B., Nickerson, J., and Owan, H. (2003). Team incentives and worker heterogeneity: An empirical analysis of the impact of teams on productivity and participation. *Journal of Political Economy*, 111(3):465–497.
- Hansen, K., Heckman, J., and Mullen, K. (2004). The effect of schooling and ability on achievement test scores. *Journal of Econometrics*, 121:39 – 98.
- Hansen, Z., Owan, H., and Pan, J. (2006). The impact of group diversity on performance and knowledge spillover – an experiment in a college classroom. Working Paper 12251, National Bureau of Economic Research.
- Hanushek, E. and Woessmann, L. (2008). The role of cognitive skills in economic development. *Journal of Economic Literature*, 46(3):607–668.
- Hartog, J., Van Praag, M., and Van Der Sluis, J. (2010). If you are so smart, why aren’t you an entrepreneur? returns to cognitive and social ability: Entrepreneurs versus employees. *Journal of Economics & Management Strategy*, 19(4):947–989.
- Heckman, J., Stixrud, J., and Urzua, S. (2006). The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. *Journal of Labor Economics*, 24(3):411–482.
- Hessels, J., Brixy, U., Naudé, W., and Gries, T. (2014). Skill variety, innovation and new business formation. *IZA Discussion Paper*.
- Hmieleski, K. and Ensley, M. (2007). A contextual examination of new venture performance: entrepreneur leadership behavior, top management team heterogeneity, and environmental dynamism. *Journal of Organizational Behavior*, 28(7):865–889.

- Hoogendoorn, S., Oosterbeek, H., and Van Praag, C. (2013). The impact of gender diversity on the performance of business teams: Evidence from a field experiment. *Management Science*, 59(7):1514–1528.
- Hoogendoorn, S., Parker, S., and Van Praag, C. (2012). Ability dispersion and team performance: A field experiment. *IZA Discussion Paper*.
- Hoogendoorn, S. and Van Praag, C. (2012). Ethnic diversity and team performance: a randomized field experiment. *IZA Discussion Paper*.
- Jovanovic, B. (1982). Selection and the evolution of industry. *Econometrica*, 50(3):649–670.
- Kahane, L., Longley, N., and Simmons, R. (2013). The effects of coworker heterogeneity on firm-level output: assessing the impacts of cultural and language diversity in the national hockey league. *Review of Economics and Statistics*, 95(1):302–314.
- Kamm, J., Shuman, J., Seeger, J., and Nurick, A. (1990). Entrepreneurial teams in new venture creation: A research agenda. *Entrepreneurship: Theory and Practice*, 14(4):7–17.
- Kilduff, M., Angelmar, R., and Mehra, A. (2000). Top management-team diversity and firm performance: Examining the role of cognitions. *Organization Science*, 11(1):21–34.
- Klotz, A., Hmieleski, K., Bradley, B., and Busenitz, L. (2014). New venture teams a review of the literature and roadmap for future research. *Journal of Management*, 40(1):226–255.
- Lazear, E. (1999). Globalization and the market for teammates. *The Economic Journal*, 109(454):15–40.
- Lazear, E. (2005). Entrepreneurship. *Journal of Labor Economics*, 23(4):pp. 649–680.
- Lechler, T. (2001). Social interaction: A determinant of entrepreneurial team venture success. *Small Business Economics*, 16(4):263–278.
- Lechmann, D. and Schnabel, C. (2011). Are the self-employed really jacks-of-all-trades? testing the assumptions and implications of lazear’s theory of entrepreneurship with german data. *IZA Discussion Paper*.
- Leonard, J., Levine, D., and Joshi, A. (2004). Do birds of a feather shop together? the effects on performance of employees’ similarity with one another and with customers. *Journal of Organizational Behavior*, 25(6):731–754.
- Lucas, R. (1978). On the size distribution of business firms. *The Bell Journal of Economics*, 9(2):508–523.
- Mäs, M., Flache, A., Takács, K., and Jehn, K. (2013). In the short term we divide, in the long term we unite: Demographic crisscrossing and the effects of faultlines on subgroup polarization. *Organization Science*, 24(3):716–736.
- Oberschachtsiek, D. (2012). The experience of the founder and self-employment duration: a comparative advantage approach. *Small Business Economics*, 39(1):1–17.

- Parker, S. (2009). The economics of entrepreneurship. *Cambridge Books*.
- Parker, S. and Van Praag, C. (2006). Schooling, capital constraints, and entrepreneurial performance: The endogenous triangle. *Journal of Business & Economic Statistics*, 24(4):416–431.
- Richard, O., Barnett, T., Dwyer, S., and Chadwick, K. (2004). Cultural diversity in management, firm performance, and the moderating role of entrepreneurial orientation dimensions. *Academy of Management Journal*, 47(2):255–266.
- Rosendahl Huber, L., Sloof, R., and van Praag, C. (2012). The effect of early entrepreneurship education: Evidence from a randomized field experiment. *Tinbergen Institute Discussion paper*, pages 1–32.
- Silva, O. (2007). The jack-of-all-trades entrepreneur: Innate talent or acquired skill? *Economics Letters*, 97(2):118–123.
- Stuetzer, M., Goethner, M., and Cantner, U. (2012). Do balanced skills help nascent entrepreneurs to make progress in the venture creation process? *Economics Letters*, 117(1):186–188.
- Stuetzer, M., Obschonka, M., and Schmitt-Rodermund, E. (2013). Balanced skills among nascent entrepreneurs. *Small Business Economics*, 41(1):93–114.
- Van Praag, C. and Cramer, J. (2001). The roots of entrepreneurship and labour demand: Individual ability and low risk aversion. *Economica*, 68(269):45–62.
- Van Praag, C., van Witteloostuijn, A., and van der Sluis, J. (2013). The higher returns to formal education for entrepreneurs versus employees. *Small Business Economics*, 40(2):375–396.
- Wagner, J. (2006). Are nascent entrepreneurs jacks-of-all-trades? a test of Lazear’s theory of entrepreneurship with German data. *Applied Economics*, 38(20):2415–2419.
- West, G. (2007). Collective cognition: When entrepreneurial teams, not individuals, make decisions. *Entrepreneurship Theory and Practice*, 31(1):77–102.

Appendix

Table A1: Random assignment of individual specialists to teams

	Verbal specialists		Math specialists	
	<i>Mixed-specialist-team dummy</i>		<i>Mixed-specialist-team dummy</i>	
Age	-0.06	(0.11)	-0.13	(0.11)
Female	0.06	(0.06)	-0.09*	(0.05)
Math score	-3.80	(2.93)	5.93*	(3.23)
Verbal score	-3.08	(3.47)	0.66	(3.55)
High school track (1-5)	-0.01	(0.19)	-0.14	(0.22)
Nationality parents: both non-Dutch	0.10	(0.09)	-0.002	(0.09)
Entrepreneurship knowledge	-0.25	(0.22)	-0.38	(0.23)
Number of teams	227		200	

*Note: The coefficients in each cell come from separate regressions of the row variable on a dummy variable that takes the value 1 if an individual was assigned to a mixed-specialist-team and 0 if an individual was assigned to a pure specialist team matching his own type. Observations are clustered at the class level, robust standard errors in parentheses. Math- and verbal score are based on the nationwide exam and the converted scores from the student tracking system, e available for 182 math specialists and 181 verbal specialists. High school track corresponds to the intended future high school track and ranges from 1 (pre-vocational secondary education) to 5 (pre-university education). Entrepreneurship knowledge is based on the number of correct answers to eight questions about entrepreneurship. */**/** indicates significance at the 10%/5%/1%-level.

Table A2: The effect of team type on team performance (without controls)

	Normalized Team Rank		Value of own shares		Money won in Tournament	
	(1)		(2)		(3)	
Math-specialist-team	0.04	(0.06)	8.26	(26.49)	0.40	(0.36)
Mixed-specialist-team	-0.11*	(0.06)	-31.03	(20.43)	-0.25	(0.26)
Verbal-specialist-team	-0.19***	(0.05)	-61.93***	(20.00)	-0.71***	(0.23)
<i>(omitted category: JAT-team)</i>						
Constant	-0.53***	(0.02)	122.31***	(15.37)	1.33***	(0.11)
Number of teams	179		113		179	
R^2	0.07		0.05		0.05	

* Note: */**/** indicates a significance at the 10%/5%/1%-level. Observations are clustered at the class level. Robust standard errors in parentheses. The results for Value of own shares were obtained using a sub sample for which this outcome measure was the only determinant for team ranking.

