IZA DP No. 7044

Ability Dispersion and Team Performance: A Field Experiment

Sander Hoogendoorn
Simon C. Parker
Mirjam van Praag

November 2012

Ability Dispersion and Team Performance: A Field Experiment

Sander Hoogendoorn
*University of Amsterdam
and Tinbergen Institute*

Simon C. Parker
*University of Western Ontario
and IZA*

Mirjam van Praag
*University of Amsterdam,
Tinbergen Institute and IZA*

IZA Discussion Paper No. 7044
November 2012

ABSTRACT

Ability Dispersion and Team Performance: A Field Experiment[*]

This paper studies the impact of diversity in cognitive ability among members of a team on their performance. We conduct a large field experiment in which teams start up and manage real companies under identical circumstances. Exogenous variation in – otherwise random – team composition is imposed by assigning individuals to teams based on their measured cognitive abilities. The setting is one of business management practices in the longer run where tasks are diverse and involve complex decision-making. We propose a model in which greater ability dispersion generates greater knowledge for a team, but also increases the costs of monitoring necessitated by moral hazard. Consistent with the predictions of our model, we find that team performance as measured in terms of sales, profits and profits per share first increases, and then decreases, with ability dispersion. Teams with a moderate degree of ability dispersion also experience fewer dismissals due to fewer shirking members in those teams.

Corresponding author:

Mirjam Van Praag
Amsterdam School of Economics
University of Amsterdam
Roetersstraat 11
1018 WB Amsterdam
The Netherlands
E-mail: c.m.vanpraag@uva.nl

1   Introduction

Few topics in economics are not associated with cognitive ability. Cognitive ability shapes individuals' behavior, decision-making and performance outcomes (Cutler and Lleras-Muney, 2010; Dohmen et al., 2010; Grinblatt et al., 2011), and is a major determinant of individual earnings, income distribution and — at the aggregate level — economic growth (Hanushek and Woessmann, 2008). Ability considerations carry over naturally from individual settings to those involving teams of individuals: studies of team performance invariably emphasize the importance of individual team member abilities (e.g., Stewart, 2006).

Decision-making in organizations is increasingly performed by teams rather than by individuals (Hamilton et al., 2003; Woolley et al., 2010). Consequently, the optimal composition of teams in terms of member abilities can be regarded as a valuable organizational asset. Yet, despite it being widely believed that (cognitive) abilities of members affect overall team performance, the precise impact of ability dispersion on team performance remains poorly understood (Hamilton et al., 2012). Specifically, we lack evidence about whether ability diversity of team members is or is not conducive to the performance of teams. Such evidence is potentially useful for managers seeking to select members of internal work teams, as well as being of interest in its own right.

This paper examines the effect of the dispersion of cognitive ability (hereafter, just 'ability') on the performance of business teams in a field experiment. We study teams of undergraduate students who are required to start up and manage a real company as a compulsory part of the curriculum in an international business program in the Netherlands. Companies are simultaneously founded on a level playing field and students face strong incentives that align their interests with the business performance of their company. Our experiment randomizes 573 students into 49 teams conditional on their measured cognitive abilities. We ensure a relatively large exogenous variation in ability dispersion between teams to help probe non-linearities in the relationship between ability dispersion and team performance.

We believe there are two principal advantages of our empirical design. First, a field experiment can establish a causal relationship between team composition and performance, in contrast to regression-based approaches in which members are free to self-select into and out of teams (Hansen et al., 2006). Second, relative to prior empirical studies which have analyzed settings involving laboratory experiments (Woolley et al., 2010), unskilled work tasks (e.g., Hamilton et al., 2003), or competitive sports (Kahn, 2000), our field experiment closely resembles the functioning of teams co-operating on a complicated real business project. The case of business (management) teams is worth understanding well. These teams often exert considerable influence on the performance of public and private organizations since they take decisions of strategic and operational importance. Moreover, the tasks of these teams are often complex and broad in scope, entailing the sustained application of members' cognitive abilities. This is precisely the context of the field experiment conducted in this paper.

Our findings can be summarized as follows. Team performance as measured in terms of sales, profits and profits per share first increases at low levels of ability dispersion up to a maximum before

decreasing at higher levels of dispersion. Controlling for the average cognitive ability of teams, performance is maximized at a coefficient of variation in cognitive ability of approximately 0.25 (the sample average is 0.22). Teams of moderate ability dispersion also experience fewer dismissals during the program, although this does not chiefly explain why those teams perform better.

Prior theorizing turns out to be unable to explain these results. Previous theories have high-lighted benefits to, or costs from, diversity — predicting linear relationships between member ability diversity and team performance. Thus, if the inputs of members are complementary in team production, performance is maximized by matching individuals of similar abilities (Kremer, 1993). However, if members' inputs are substitutable, heterogeneous teams in terms of ability maximize performance (Prat, 2002).[1] These arguments have been applied to explain performance in team sports, for example, Gould and Winter (2009) in the case of Major League Baseball, and Franck and Nüesch (2010) in the case of German professional soccer. Other reasons why teams might benefit from heterogeneity in ability among their members include the imposition of valuable high team production norms by a few powerful able members and learning by less able from more able team members (Hamilton et al., 2003).[2]

In order to understand our results, we develop a novel model in which intra-team ability dispersion has both beneficial and detrimental effects. Low levels of ability dispersion in the model are associated with smaller pools of knowledge and, hence, team outcomes of limited scope and value. Greater dispersion is associated with larger pools of knowledge that produce more valuable team output, but this comes at an increasing cost. Team members can choose to shirk rather than provide effort, creating a free-riding incentive. Teams discourage shirking by monitoring their members; but the success of the monitoring technology is stochastic, decreasing with team member diversity.[3] At sufficiently high levels of ability dispersion, the probability of being caught and dismissed for shirking declines so much that shirking reduces team performance. Hence, moderate levels of ability dispersion in a team are associated with maximal team performance, i.e., the relationship between team performance and ability dispersion exhibits an inverse U-shaped pattern.

The remainder of this paper is organized as follows. Section 2 presents the theoretical model. Section 3 outlines the context, design and data of our field experiment. Section 4 presents the empirical findings. Section 5 discusses and concludes.

---

[1] In a model of conjunctive team production the lowest-ability member (i.e., the weakest link) in the extreme case determines team performance, whereas a model of disjunctive team production implies that ultimately the highest-ability member (i.e., the superstar) determines the performance of teams (see also Hong and Page, 2001, 2004; Iranzo et al., 2008; Rosen, 1981).

[2] Consistent with the predictions of their model, Hamilton et al. (2003) find that teams' ability dispersion positively affects their output using high-frequency productivity data from a garment plant.

[3] A related deterrence mechanism is peer monitoring. For an analysis of peer effects in the workplace, see Falk and Ichino (2006), Kandel and Lazear (1992), and Mas and Moretti (2009).

2 Model

2.1 Set up

We commence by analyzing a two-member team, in which both members create value by combining their knowledge. Members have potentially different cognitive abilities, $x_i \in X$ and $x_j \in X$, where X is the set of cognitive abilities in the general population. We have $\min\{X\} = x_L$ and $\max\{X\} = x_U > x_L$. Denote the average cognitive ability of the team by $\bar{x} = (x_i + x_j)/2$ and the difference between the two cognitive abilities by $d = |x_j - x_i|$: $d \in [0, x_U - x_L]$. There is a large number of teams in the population, whose total mass is unity. Decisions of one team do not affect other teams. Different teams may have different values of $(\bar{x}, d)$; our analysis will focus on only a single team.

People of given cognitive ability also have a given set of knowledge, denoted by $\{k\} \subset K$, where K is the set of knowledge in the general population. We assume that cognitive ability x and knowledge set $\{k\}$ are related by a *correspondence*, whereby each element x of X is related via a map g to a given subset $P(K)$ of K. Formally, $g : X \to P(K)$, where $P(K)$ is known as the *power set* of K. To crystallize ideas in what follows we will work with the particular case where the power set associated with x is $\{x \pm \frac{1}{2}\Delta\}$, where $0 < \Delta < x_U - x_L$. Hence, $K = [x_L - \Delta/2, x_U + \Delta/2]$. In this case, a unique and equally sized set of knowledge, spanning Δ in size, is associated with each unique level of cognitive ability. For now, Δ will be taken to be invariant to x. Later on, we will relax this assumption and allow $\Delta = \Delta(x)$, with $\Delta'(x) > 0$, to encompass the possibility that more able people have larger knowledge sets than less able people.

Note that team members of low and high cognitive ability both have productive roles in this set-up. That is because both members possess unique knowledge. For example, a very able team member might possess detailed scientific knowledge about an invention, which the less able team member may lack. Yet, the less able member might possess knowledge about, e.g., salesmanship or market conditions, which the more able member lacks. Team members pool their knowledge: team value, v, therefore depends on the union of their (non-overlapping) knowledge sets. Thus, the value of a team is greater than that of a single team member, as long as $x_j \neq x_i$, i.e., as long as $d > 0$.

Baumol and Strom (2010) describe how Matthew Boulton's knowledge of eighteenth century industrial market needs complemented James Watt's technical knowledge of steam engines and crank technology, forging a partnership that pioneered and disseminated the rotary-motion engines which powered the Industrial Revolution. As this example shows, the union of disparate knowledge generates value without either partner needing to acquire the knowledge possessed by the other (see also Lazear, 1999).

If x_i and x_j are very similar, knowledge spans will overlap and the gains from being in a team are modest. Hence, firm value is greater the larger the ability difference d as long as $d < \Delta$. Once $d \geq \Delta$, firm value is no longer increasing in d since maximal knowledge pooling has occurred. Hence,

the firm's value is increasing in v, where

$$v = v(d) = \begin{cases} \Delta + d & \text{for } d < \Delta \\ 2\Delta & \text{for } d \geq \Delta \end{cases} \tag{1}$$

See Figure 1 for an illustration of three different cases corresponding to (1). Note that none of the analysis or results that follow depend on the discontinuity of $v(d)$ in (1): a smooth function for $v(d)$ which exhibits diminishing marginal returns to d would generate the same qualitative results.

Total firm value depends not only on $v(d)$ but also on the efforts of both team members. Effort is privately costly for each member but is non-contractible and observable only by the member who exerts it: it cannot be credibly communicated or signaled to the co-member (e.g., Laffont and Martimort, 2002). Each of the member efforts are inputs into the production process. Each effort, denoted by e_i and e_j for the two members, either takes a value of one (full effort, which incurs private and non-publicly-observed idiosyncratic effort costs of $c_i > 0$ and $c_j > 0$ respectively, where $c_i \neq c_j$ in general) or zero ('shirking': no effort, so no effort cost). Effort costs are also private information, which cannot be credibly communicated or signaled to the co-member or a third party. The density function of effort costs in the population is $\gamma(c) > 0$ for $c \geq 0$; the cumulative distribution function is $\Gamma(c) = \int_0^c \gamma(\chi) \, d\chi \in [0, 1]$.

Each team member chooses their effort given expectations (derived below) of the other member's effort. Total gross firm value is

$$\tilde{V} = v(d)[f(e_i, e_j) + \tilde{\omega}] \tag{2}$$

where f is an increasing function of both arguments. Without any important loss of generality, we will assume that $f(e_i, e_j) = e_i + e_j$. Hence, $f(0,0) = 0$: zero effort by both members yields zero expected firm value. The separability of the production function helps keep the decision-making separable at the individual level, which will enable us to study a game in pure strategies with a unique equilibrium. Let ω denote the outcome of a mean-zero random variable $\tilde{\omega}$ (e.g., this could capture stochastic demand for the firm's output). Only the joint return $e_i + e_j + \omega$ is observed. So while a member knows her own effort with certainty, she cannot infer the other effort once V is observed, since ω is also unobserved.

This is important because it means that monitoring is the only way that members can measure the effort of co-members. At the outset, both members agree on the following contract: (a) realized firm value is equally shared at the end of the period if neither member is dismissed, and (b) any member who is discovered via monitoring to have supplied zero effort is dismissed from the team. They receive zero output and pay a penalty of $\sigma > 0$. The remaining member takes the remaining output in its entirety. The reason for (a) is that costs and efforts are unobserved and output is a joint product: hence, an alternative *ex ante* compensation scheme based on effort-related equity shares is not feasible. For (b), dismissal in response to zero effort by the other member is optimal for the effort-providing member while providing an *ex ante* incentive for both members to exert costly effort. The penalty σ paid by a dismissed member could be lost capital or lost reputation,

5

(a)
$v(d) = 2\Delta$

X_i X_j

$X_i - \Delta/2$ $X_i + \Delta/2$ $X_j - \Delta/2$ $X_j + \Delta/2$

(b)
$v(d) = 3/2\Delta$

X_i X_j

$\Delta/2$

$X_i - \Delta/2$ $X_j + \Delta/2$

(c)
$v(d) = \Delta$

$X_i = X_j$

$X_i - \Delta/2$ $X_j + \Delta/2$
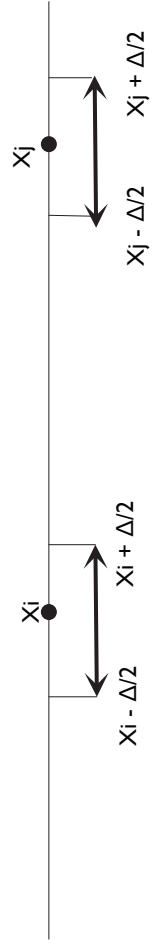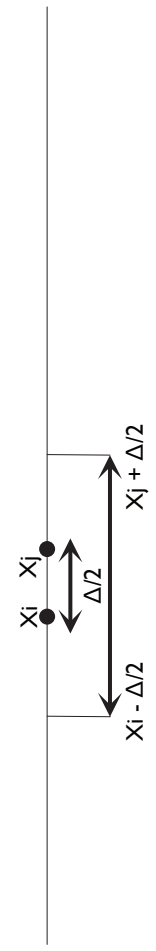
Figure 1. Illustration of three different cases corresponding to (1)

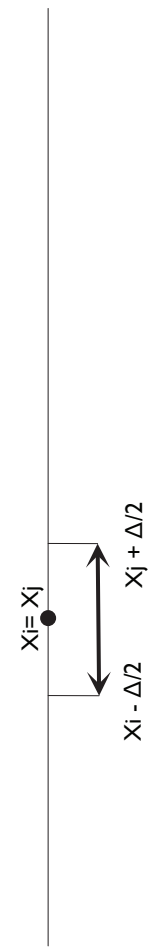which detracts from their future expected economic prospects. The probability of dismissal will be derived endogenously below.

With probability $p(d)$, monitoring reveals to a member any shirking by her co-member. Here, p is assumed to be a smooth and continuous function. Detection of shirking is assumed to be more difficult in teams where there is greater distance d between the members, reflecting the difficulty of disentangling effort from the productivity of that effort.[4] Hence, $p'(d) < 0$, with $p(0) = 1$, $p(\infty) = 0$ and $p''(d) > 0$. We assume that the monitoring technology never falsely indicates shirking when effort was actually supplied. Monitoring is taken to be effortless and costless without loss of generality.

We do not study the determinants of team formation, taking (in accordance with our empirical data) the composition of teams to be exogenous. It is worth briefly underlining the importance of assuming exogenous team composition. We want to predict how individual and team efforts, performance and dismissals vary with different amounts of team diversity. If instead we allowed for endogenous self-selection of diverse individuals into teams, we would only be able to analyze these relationships for that subset of team diversities associated with some given (assumed) self-selection process. We are interested in analyzing what could happen under a range of alternative sorting arrangements. Furthermore, the assumption of exogenous team assignment fits directly with our experimental design, which can therefore provide causal evidence about the consequences of a range of different team diversities.

2.2 Payoffs and optimal effort choices

Both members know d, but are symmetrically uninformed about the private effort costs faced by their co-members and, hence, the effort that their co-members will supply. They each believe with probability θ (derived below) that their co-member will exert high-effort and with probability $1 - \theta$ that they will shirk. Expected payoffs under low and high effort by member j are

$$E(R_j | e_j = 0) = \theta v(d)(1 - p(d))/2 - p(d)\sigma \tag{3}$$

$$E(R_j | e_j = 1) = v(d)\left[\theta + \frac{(1 - \theta)(1 - p(d))}{2} + (1 - \theta)p(d)\right] - c_j \tag{4}$$

Evidently, j's optimal effort choice is given by

$$e_j^* = \begin{cases} 0 & \text{if } E(R_j | e_j = 0) \geq E(R_j | e_j = 1) \\ 1 & \text{otherwise} \end{cases} \tag{5}$$

---

[4]Thus, taking a hypothetical extension of the Boulton and Watt example, Watt might be unsure about whether any initial lack of success by Boulton in selling engines is attributable to limited sales effort by Boulton or rather full but ineffective sales effort. Likewise, Boulton might be unsure whether any initial lack of success by Watt in building a new engine is attributable to limited technical effort by Watt or rather full but ineffective technical effort. Our assumption of $p'(d) < 0$ is consistent with the idea that had Boulton and Watt been more dissimilar, their difficulties of disentangling lack of effort from ineffective deployment of effort would have been even greater.

After rearranging and collecting terms, the condition for $e_j^* = 1$ is

$$c_j < v(d)(1 + p(d))/2 + p(d)\sigma \tag{6}$$

We assume that the equilibrium strategy is the same for both members, i.e., an analogous expression to (6) exists for i (replace c_j by c_i on the LHS). Hence, we have a symmetric problem.

Member j knows her own c_j but not c_i; i knows c_i but not c_j. While j does not know c_i she does know the distribution of types, so her subjective probability that the inequality in (6) holds is

$$\theta = \theta\big(v(d), p(d), \sigma\big) = \Gamma\left[v(d)(1 + p(d))/2 + p(d)\sigma\right] \tag{7}$$

Differentiate (7) to obtain:

$$\frac{\partial \theta}{\partial v(d)} = \frac{\gamma(\cdot)(1 + p(d))}{2} > 0 \tag{8}$$

$$\frac{\partial \theta}{\partial p(d)} = \gamma(\cdot)\left(\frac{v(d)}{2} + \sigma\right) > 0 \tag{9}$$

The signs of these derivatives both make intuitive sense: a greater return from effort increases the probability that effort is exerted, as does a greater probability of being caught if one shirks.

Note that θ is fixed for a given team (i.e., for a given d), though its value varies across teams with different values of d. Because the researcher does not observe each private effort choice either, we can state our first proposition in terms of expected performance across different teams.

Proposition 1. *Provided the effectiveness of monitoring does not decline too rapidly as members become infinitesimally diverse, i.e., provided that*

$$|p'(0)| < \frac{2[\Gamma(\Delta + \sigma) + \Delta\gamma(\Delta + \sigma)]}{\gamma(\Delta + \sigma)(\sigma + \Delta/2)} \tag{10}$$

then expected team performance exhibits an inverse U-shaped relationship with the diversity of team member ability.

Proof. Both members provide effort with probability θ^2, while only one member provides effort with probability $\binom{2}{1}\theta(1 - \theta)$. Hence, expected firm performance and its derivative with respect to d, are

$$E(\tilde{V}) = 2v(d)\big(\theta^2 + \theta(1 - \theta)\big) = 2v(d)\theta$$

$$\frac{\partial E(\tilde{V})}{\partial d} = 2\theta v'(d) + 2v(d)\left(\frac{\partial \theta}{\partial v(d)}v'(d) + \frac{\partial \theta}{\partial p(d)}p'(d)\right) \tag{11}$$

As shown above, the two derivatives (8) and (9) which appear in (11) in large brackets are both positive. Also $v'(d)$ is positive for $d < \Delta$ and zero for $d > \Delta$, while $p'(d) < 0 \; \forall d$. Hence, the overall

8

derivative (11) is certainly positive at $d = 0$ if the condition in the body of the proposition holds, and remains positive as long as $d < \Delta$ since $p''(d) > 0$. But for $d > \Delta$, $v'(d) = 0$ so by inspection (11) turns negative. Hence, expected firm performance $E(\tilde{V})$ is an inverse U-shaped function of d.

The intuition for Proposition 1 is as follows. Low levels of team diversity $(0 \leq d < \Delta)$ are associated with smaller pools of knowledge and, hence, team outcomes of limited scope and value. Higher levels of dispersion are associated with larger pools of knowledge that produce more valuable team output. However, there is a limit to this benefit of greater diversity; at sufficiently high levels of dispersion the probability of being caught and dismissed for shirking induces such low member effort that teams lose productivity and firm performance worsens.

The role of the condition in the body of Proposition 1 has a straightforward interpretation. If monitoring effectiveness declines very rapidly as members become marginally different from each other, then effort declines so rapidly that the ensuing decline in performance dominates any positive effect from greater knowledge pooling. Thus, if the condition in the proposition does not hold, firm performance will be strictly declining in d.

2.3 Dismissals

Next we ask which teams are most prone to dismissals of one member. In what follows, we will not pay attention to cases where both members end up dismissing each other (a case referred to as 'team collapse'). Indeed, there are no team collapses in the data to motivate such an analysis here. We will instead focus on cases where only one member is dismissed.

A dismissal occurs if one member shirks and is caught while the co-member either does not shirk (or does and is not caught; if both members are caught shirking there is a team collapse instead). There are two ways the event 'one member caught shirking' can occur, so the probability of a dismissal, $\Psi(d, \theta)$, is

$$\Psi(d, \theta) = 2p(d)(1 - \theta)[\theta + (1 - \theta)(1 - p(d))] \tag{12}$$

The derivative with respect to d is

$$\frac{\partial \Psi(d, \theta)}{\partial d} = 2(1 - \theta)\Upsilon(d)p'(d) - 2p(d)\Upsilon(d)\left(\frac{\partial \theta}{\partial v(d)}v'(d) + \frac{\partial \theta}{\partial p(d)}p'(d)\right) \tag{13}$$

where $\Upsilon(d) = 1 - 2p(d)(1 - \theta)$. If (as is supported by the evidence in our study) less than half of all teams contain at least one detected shirker, then $p(d)(1 - \theta) < 0.5$, whence $\Upsilon(d) > 0 \ \forall d$. A sufficient condition for this to hold is $2\Gamma(\Delta + \sigma) > 1$, which we will assume for the next proposition.

Proposition 2. *A sufficient condition for the incidence of dismissals to be a strictly decreasing function of diversity, d, is*

$$p(d)\frac{\partial \theta}{\partial p(d)} < 1 - \theta \quad \forall d \tag{14}$$

9

If the inequality in (14) is reversed, and if in addition

$$|p'(0)| < \frac{\gamma(\Delta + \sigma)}{1 - \Gamma(\Delta + \sigma) - \gamma(\Delta + \sigma)(\sigma + \Delta/2)} \qquad (15)$$

then the incidence of dismissals is a U-shaped function of d.

Proof. Rearrange (13) to obtain

$$\frac{\partial \Psi(d, \theta)}{\partial d} = 2\Upsilon(d)p'(d)\left[1 - \theta - p(d)\frac{\partial \theta}{\partial p(d)}\right] - 2p(d)\Upsilon(d)\frac{\partial \theta}{\partial v(d)}v'(d) \qquad (16)$$

If (14) holds, then (16) is strictly negative. If (14) does not hold while (15) does, then the final negative term of (16) dominates at $d = 0$, implying Ψ initially declines in d. For $d > \Delta$ we have $v'(d) = 0$ and the first term of (16) remains. This is positive if the inequality in (14) is reversed, proving the result.

The two possible dismissal profiles outlined in Proposition 2 depend on the sensitivity of effort to the probability of detecting shirkers. If effort is relatively insensitive to the probability that shirkers are detected (i.e., if (14) holds) then the likelihood of dismissals decreases as knowledge pooling generates benefits from effort (for low to moderate dispersion) and as the effectiveness at catching shirkers decreases (at high diversity). On the other hand, if effort is highly sensitive to the probability that shirkers are detected (i.e., if the inequality in (14) is reversed), and if the probability of detecting shirkers does not decrease too rapidly in response to a marginal increase in dispersion (condition (15)), then teams with low diversity supply more effort in response to benefits from knowledge pooling — leading to fewer dismissals initially. However, in more diverse teams where the marginal benefits of pooling have attenuated, members supply so much less effort that, even though the probability of being detected has decreased, there are more culprits to catch and hence the incidence of dismissals rises.

2.4   A simple extension

The model can be extended to treat the case with more than two team members. A separate appendix, available on request, shows that it is possible to extend the analysis to this case, such that Propositions 1 and 2 continue to apply. An analytically simpler, but more far-reaching, extension allows cognitive ability to affect the size of knowledge sets. This possibility introduces a novel prediction, summarized in a new Proposition 3, below.

To model this extension, suppose $\Delta'(x) > 0$ so if $x_j > x_i$ then $\Delta(x_j) > \Delta(x_i)$. Then the

following equation replaces (1):

$$v(d) = v(d, \bar{x}) = \begin{cases} \Delta(x_j) & \text{for } d < [\Delta(x_j) - \Delta(x_i)]/2 \\ d + [\Delta(x_j) + \Delta(x_i)]/2 & \text{for } [\Delta(x_j) - \Delta(x_i)]/2 \leq d < [\Delta(x_j) + \Delta(x_i)]/2 \\ \Delta(x_j) + \Delta(x_i) & \text{for } d \geq [\Delta(x_j) + \Delta(x_i)]/2 \end{cases}$$
(17)

Teams' average cognitive ability $\bar{x}$ qualifies as a valid argument of (17) because $\bar{x}$ is associated with greater x_j and/or x_i, which in all cases increases v. Taking this into account and using (2), the next proposition follows immediately.

Proposition 3. *If knowledge sets and cognitive ability are positively related, and the condition* (10) *of Proposition 1 holds, then*
(a) Expected team performance exhibits first a declining and then an inverse U-shaped relationship with the diversity of team member ability, and
(b) Expected team performance is positively related to average team ability.

Note that there are two principal differences between Proposition 3 and Proposition 1. First, the flat initial section of the $v(d)$ function in (17) gives rise to an initial declining segment of the effort and performance relationships with respect to d, before the inverse U-segments emerge. A testable implication of this prediction is that a third-order polynomial function is needed to represent the performance-dispersion relationship if knowledge breadth really does increase with cognitive ability. In contrast, a third-order term will be insignificantly different from zero if knowledge breadth is invariant to cognitive ability, as assumed to be the case in Proposition 1.

Second, Proposition 3 predicts the importance of including an additional explanatory variable in a model of team performance if knowledge breadth increases with cognitive ability: namely, the average cognitive ability of team members. In contrast, this explanatory variable will be insignificantly different from zero if knowledge breadth is invariant to cognitive ability, as assumed to be the case in Proposition 1. Like the significance of the higher order term, this too is an easily testable prediction.

3   Context, design and data

3.1   *Context*

The teams in our field experiment are teams of undergraduate students that have to start up and manage a real company as a compulsory part of the curriculum at the department of international business studies of the Amsterdam College of Applied Sciences.[5] The entrepreneurship program

---

[5]The department of international business studies at the Amsterdam College of Applied Sciences is subdivided into five sub-departments/fields of study: business management, management, trade management Asia, business languages and financial management. This division is relevant for our study since the assignment of students to teams takes place within these sub-departments/fields of study (discussed below).

covers about one-fifth of students' first-year undergraduate curriculum. This program is organized in collaboration with Junior Achievement (JA), which is the worldwide leading provider of educational programs in entrepreneurship (Oosterbeek et al., 2010). Companies in the entrepreneurship program are simultaneously founded on a level playing field and dissolved after one academic year. The experiment was performed in 2009-2010 and its (description of the) context and incentives are similar to the experiments reported in Hoogendoorn and Van Praag (2012) and Hoogendoorn et al. (2013).

During the program students have to: raise capital by issuing shares; appoint officers and delegate tasks; produce and market products or services; keep the accounts; and conduct shareholders' meetings. Hence, students execute a substantial and genuinely joint task that requires them to establish roles, build up relationships, and create routines and processes in order to maximize shareholder value. Moreover, students face strong incentives that align their interests with the business performance of the company (see subsection 3.2). Each company reports to their randomly assigned professor and business coach on a regular basis. Everything about the company is real, including tax and social security payments. The program is not a business simulation. In sum, students in our experiment have to coordinate on a broad array of complex decision-making tasks that entail the sustained application of their cognitive abilities.

Companies typically proceed as follows. They start with brainstorming about potential business activities and conducting market research to select the most viable idea. There are no restrictions on the type of business activity that can be chosen. Simultaneously, teams appoint about half of their members to management positions (such as the CEO and CFO) and the other half to non-management positions, where management positions are redistributed among the non-managing part of the team halfway the program.[6] Companies further develop their chosen idea by writing a business plan, and they start raising capital by issuing shares. Other sources of financing such as personal or outside loans are not allowed. Once the business plan is authorized by the majority of shareholders at the first shareholders' meeting, business operations of teams boil down to production and marketing of the chosen products or services. All companies are dissolved at the end of the program and each team has to write an annual report that needs approval at the final shareholders' meeting. Any profits are divided among the shareholders.

Our experiment randomizes 573 students into (a by the college predefined number of) 49 teams conditional on their score in the 20-minute timed version of Raven's advanced progressive matrices test as a proxy for cognitive ability (Hamel and Schmittmann, 2006; Raven et al., 1998). The average team size is equal to approximately 12 students. Table A1 in the appendix lists the key characteristics of all 49 teams, including the product or service they market.

Cognitive ability
Raven's advanced progressive matrices are extensively used to differentiate between people of higher

---

[6]The ability composition of the entire team closely resembles that of the management team, which is possibly the more influential part of a team. Point estimates from a regression of the ability composition of the management team on the ability composition of the entire team are not significantly different from 1 before and after roles are switched.

cognitive ability (Bors and Stokes, 1998; Mills et al., 1993; Raven et al., 1993). The test requires subjects to select the missing figure out of eight possibilities that completes a logical pattern (see Figure A1 in the appendix for an example). Patterns become increasingly difficult as subjects progress. Over the past decades, Raven's advanced progressive matrices have been shown to associate with cognitive ability or intelligence in various ways. Elaborating on Spearman's notion of general cognitive ability (1927), Raven's advanced progressive matrices are found to measure fluid intelligence (Cattell, 1963), analytic intelligence (Carpenter et al., 1990), and intellectual efficiency if administered with a time limit (Hamel and Schmittmann, 2006).[7] As such, test scores on Raven's advanced progressive matrices can be interpreted as a proxy for cognitive ability. Indeed, the correlation between these test scores and students' grade point average (GPA) shows a significant and positive relationship in our sample.

Table 1. Descriptive statistics of cognitive ability

A: Cognitive ability	Mean	SD	Min	Max				
Average ability	18.60	2.53	14.00	23.22				
Ability dispersion (CV)	0.22	0.09	0.07	0.47				
B: Field of study	Stud.	Teams	Average ability			Ability dispersion		
			Mean	Min	Max	Mean	Min	Max
Business management	265	21	18.74	14.07	22.40	0.22	0.07	0.47
Management	45	4	18.67	16.27	22.22	0.23	0.18	0.33
Trade management Asia	108	10	18.92	14.78	23.22	0.20	0.10	0.35
Business languages	123	12	17.94	14.00	21.86	0.21	0.08	0.36
Financial management	32	2	19.33	18.83	19.83	0.27	0.21	0.32
Total	573	49	18.60	14.00	23.22	0.22	0.07	0.47

Note: Average and CV of ability reflect at the team level respectively average score and coefficient of variation in scores on Raven's advanced progressive matrices test.

Table 1 provides descriptive statistics of cognitive ability at the team level (panel A) and by field of study (panel B). Panel A indicates that the average number of figures correctly solved in the 20-minute timed version of Raven's advanced progressive matrices test is 18.60 out of 36 figures at maximum. We do not transform these test scores into an intelligence quotient, because this would require an additional assumption about the proper norm for first-year college students (Hamel and Schmittmann, 2006).[8] Moreover, our main interest is in the exogenous variation in cognitive

---

[7]Spearman (1927) decomposed general cognitive ability (g) into an eductive and a reproductive component, where (i) eductive ability reflects "the ability to make meaning out of confusion, the ability to generate high-level, usually nonverbal, schemata which make it easy to handle complexity", and (ii) reproductive ability reflects "the ability to absorb, recall, and reproduce information that has been made explicit and communicated from one person to another" (Raven, 2000, p. 2). Fluid intelligence, analytic intelligence, intellectual efficiency and, hence, the 20-minute timed version of Raven's advanced progressive matrices test, mainly relate to eductive ability. Nevertheless, scores on cognitive tests such as Raven's advanced progressive matrices may differ across time, gender and culture (see Calvin et al., 2011; Irwing and Lynn, 2005; Nisbett et al., 2012; Raven, 2000; Rushton and Jensen, 2005).

[8]In line with Bors and Stokes (1998) and Raven et al. (1998) we also exclude students with a test score $\leqslant 6$ from the sample. Including this group of 10 students in total may incorrectly inflate teams' ability dispersion since a test score $\leqslant 6$, in more convential intelligence terms, roughly corresponds with the cognitive ability level of elementary

ability rather than the exact level of test scores. Consistent with recent empirical studies involving professional sports (Franck and Nüesch, 2010; Papps et al., 2011), we use the coefficient of variation in test scores as a scale-invariant measure for ability dispersion in teams. Teams' coefficient of variation in test scores varies between 0.07 and 0.47 with a sample average of approximately 0.22. Panel B shows the numbers of students and teams by field of study. It also indicates that (the range of) average ability and ability dispersion of teams are similar across fields of study; possibly except for the field of financial management which accommodates only two teams.

3.2 Design

The cognitive ability of students and their background characteristics were administered one week before the start of the entrepreneurship program.[9] As outside researchers we then manipulated the ability composition of teams and randomly assigned students to teams in accordance with our imposed variation in cognitive ability. In practice, we proceeded as follows.

Within fields of study, students were divided into four quartiles per class on the basis of their test score, where 1 reflects the best quartile and 4 the worst quartile. Each class was then split up in two teams, which received either treatment A or B. Treatment A combines cognitive ability quartiles 1+2 and 3+4, and treatment B combines cognitive ability quartiles 1+4 and 2+3 in a class. Hence, '1+2 teams' have a high average ability and a low ability dispersion, '3+4 teams' have a low average ability and a low ability dispersion, '1+4 teams' have a medium average ability and a high ability dispersion, and '2+3 teams' have a medium average ability and a low ability dispersion. The assignment of students was implemented one week later by the program coordinators, who were informed about the character of our field experiment.[10] Students and business coaches were uninformed, whereas professors only knew that a research project was performed that prohibited students to switch teams. Only 6 out of 573 students managed to switch teams during the program.

Figure 2 shows frequency distributions of scores on Raven's advanced progressive matrices test at the individual and team level (average ability and ability dispersion). At the individual level test scores range from 7 to 32 figures correctly solved. The average ability of teams varies between test scores of 14 and 23, while as mentioned before teams' ability dispersion ranges from 0.07 to a coefficient of variation in test scores of 0.47. We exploit this substantial and exogenous variation in cognitive ability to study the impact of teams' heterogeneity in cognitive ability.

One might worry that the effect of ability dispersion on team performance is biased since teams of low or high average ability, by construction, tend to have a lower ability dispersion (relative to teams of medium average ability). The scatter plot of teams' average ability and ability dispersion,

---

school dropouts (which is highly unlikely for first-year college students). Students with a test score $\leqslant 6$ most likely just did not put in effort or choked while taking the test. T-tests acknowledge that those students are not significantly different from students with a test score > 6 in terms of age, gender, risk aversion and GPA.

[9]Students were kept uninformed about their score in the 20-minute timed version of Raven's advanced progressive matrices test. We presented the fact that they were tested as a standard procedure of the introductory week at their new college.

[10]A few late applicants were randomly distributed among the existing teams whereas a few 'no shows' were also randomly distributed across teams (as they did not know to which team they were assigned to at that stage).
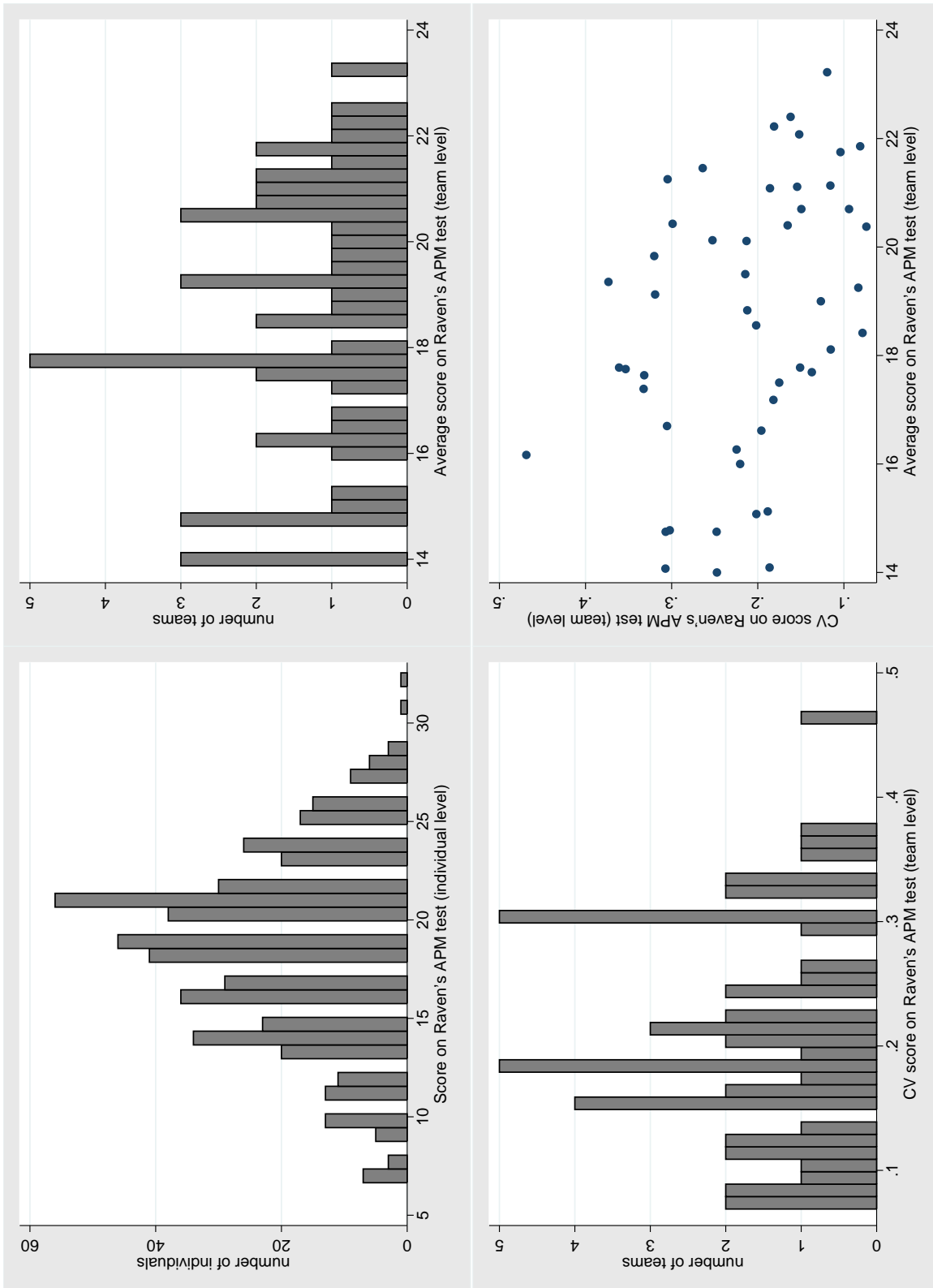
Figure 2. Frequencies and scatter of scores on Raven's advanced progressive matrices test (individual and team level)

however, does not reveal a systematic pattern that may confound a causal interpretation of the effect of ability dispersion on the business performance of teams (see Figure 2). Moreover, the results in section 4 are similar if we include only 31 medium-ability teams with average test scores not more than one standard deviation away from the average test score in the sample, i.e., with average test scores in the range of 18.60 ± 2.53 (see Table 1).

Dropout rates for first-year students in Dutch higher vocational schools, where the admission of students based on grades or previous achievements is prohibited, are on average 30% including students that switch study and/or school (ref. HBO-raad, 2010). The design of our experiment could be compromised if dropouts change the ability composition of teams. During the entrepreneurship program approximately 14% of the students dropped out or were dismissed, which reduced the average team size from 12 to 10 students.[11] Nevertheless, this did not considerably change teams' overall average ability and ability dispersion. The average ability of teams increased from a test score of 18.60 to 18.65, while teams' ability dispersion remained the same at a coefficient of variation in test scores of 0.22. The correlations of teams' average ability and ability dispersion at the start and at the end of the program are 0.96 for both measures. Students in teams of different ability composition also did not drop out more or less often (dismissals will be discussed below). In sum, we are confident that these composition dynamics did not compromise the design of our experiment.

Incentives

There are various strong incentives in place that align the interests of students with the business performance of their company. First and foremost, students can be dismissed in case of repetitive free-riding.[12] Dismissal of team members requires a two-third voting majority in the team together with the consent of the professor. It is a credible threat since the average number of dismissals is 0.35 per team and nearly 30% of the teams experiences at least one dismissal during the entrepreneurship program. Dismissal has severe consequences. Students are excluded from the program, lose its corresponding 12 credit points (out of 60 credit points in the first year) and endanger their prospect of obtaining an undergraduate degree (for which a minimum number of 45 credit points in the first year is mandatory). In section 4 we will examine to what extent dismissals vary across teams of different ability composition.

Another incentive is provided by the grade students obtain for the program from their professor, which has a substantial weight of 20% in their (first-year) grade point average. Both individual and team performance determine the program grade and their weight in the total program grade is about 50/50 (assessment of both components is based on the professor's subjective evaluation). Individual performance of students mainly entails active participation and development of competencies such as cooperation, entrepreneurial behavior and professionalism. An indicator of the effect of individual performance is the considerable average difference between the highest and lowest program grade

---

[11]Lower dropout rates than the national average at the department of international business studies of the Amsterdam College of Applied Sciences can be explained by the fact that international programs generally attract more motivated students.

[12]Interviews with program coordinators acknowledge that the main cause of dismissals is shirking of team members.

within a team of 1.5 on a 10-points scale (s.d. 0.8). The relevance of team performance for the program grade is indicated by the significant and positive correlations between teams' average program grade and business outcomes.

Virtually all students own one or more shares in their companies (with a nominal value of 20 euros per share). Roughly half of the shares are owned by team members themselves; the remaining shareholders are usually family members, friends and/or acquaintances. The mean number of shares issued is 57 (s.d. 26.1), while the minimum and maximum numbers of shares sold are respectively 21 and 135. Finally, teams participate in a formal business competition. Six selected teams present their results in a 'business pitch' at the end of the program to a jury of entrepreneurs who choose a winner based on business outcomes and presentations. The winning team obtains a cup, often gets some (local) press attention and represents the college in a national competition. All in all, the incentives discussed above ensure that students care about the business performance of their company.

3.3 Data

We accessed various data sources to collect information about individuals and teams. One week before the start of the entrepreneurship program students took the 20-minute timed version of Raven's advanced progressive matrices test and filled out a pretreatment questionnaire that mainly covered their background characteristics (response rate: 89%). Simultaneously, we received administrative data to assist us in assigning students to teams. At the end of the program, students filled out a posttreatment questionnaire that queried team characteristics and processes (response rate: 68%). We then also obtained the approved annual reports, which contain information about the business performance of teams. The data that we collected were used to: construct exogenous variation in cognitive ability across teams (see subsection 3.2); test the predictions of our model (see section 4); and assess whether the assignment of students to teams was random conditional on their cognitive ability (discussed in this subsection).

Table 2 provides descriptive statistics of individuals and teams. It shows that students are 20.9 years old on average and 43% of them are female. We also measure risk aversion (Dohmen et al.,

Table 2. Descriptive statistics of individual and team characteristics

	Scale	Mean	SD	Min	Max
Individual level					
Age	years	20.93	2.09	17.89	32.86
Gender (dummy = 1 if female)	0/1	0.43	0.50	0.00	1.00
Risk aversion	1-11	7.45	2.00	1.00	11.00
Grade point average	1-10	6.38	0.28	5.90	6.98
Team level					
Size (at baseline)	persons	11.69	2.27	7.00	17.00
Dismissals (number)	number	0.35	0.60	0.00	2.00
Dismissals (incidence)	0/1	0.29	0.46	0.00	1.00

2011) and students' scholastic achievements just before entering the college (indicated by 'grade point average'). At the team level, Table 2 indicates considerable variation in the number and incidence of dismissals across teams. Section 4 discusses whether dismissals vary in relation to the ability composition of teams.

Business performance is operationalized by four measures: sales, profits, a binary indicator for positive profits and profits per share. We include a binary indicator for positive profits to account for the fact that students tend to view as the bottom line result whether or not they are able to satisfy shareholders. Table 3 shows that average sales for all 49 teams are equal to 902 euros and that profits are 24 euros on average. More than half of the teams makes a profit (57%) and average profits per share amount to 0.62 euros. All three profit measures are significantly and positively correlated with sales.

If we split the sample into teams of low (mean<17), moderate ($17\geq$mean≤21) and high (mean>21) average ability, descriptive statistics suggest that teams of moderate average ability perform slightly better on the different business outcomes. However, note that these descriptives are very sensitive to the exact location of particularly the second cutoff point. A cutoff at a mean test score of 20 (instead of 21), for example, would already imply that teams of high average ability achieve better results. If we split up the sample in teams of low (CV<0.15), moderate ($0.15\geq$CV≤0.30) and high (CV>0.30) ability dispersion, teams of moderate ability dispersion tend to have higher sales, profits and profits per share than teams in the other two categories, on average. This ranking is rather insensitive to the precise location of the cutoff points. The effect of the average level and dispersion of ability in teams will be examined more formally in section 4.

Randomization

To assess whether the assignment of students to teams was truly random (conditional on their cognitive ability), we regress background characteristics of low ability and high ability students on the average test score in their team, the team's coefficient of variation in test scores and its square. This is consistent with the team level specifications of the main results in section 4.

Panel A1 of Table 4 shows that background characteristics of low ability students do not systematically vary across teams of different ability composition. Hence, low ability students in teams of low ability dispersion are not significantly different from low ability students in teams of high ability dispersion. The same holds for background characteristics of high ability students (see panel A2). Low ability and high ability students assigned to teams of distinct ability composition are also not more or less likely to follow a specific field of study (not tabulated).

In a similar fashion, panel B of Table 4 examines at the team level whether (average) background characteristics of students correlate with the ability composition of teams. Again, there are no systematic differences between teams of different ability composition. Note that the first-year students in our sample have roughly similar capacities since they are all relatively young and all study at the department of international business studies of the Amsterdam College of Applied Sciences.[13]

---

[13]Since the randomization checks in this subsection fail to find any pretreatment differences (that may contaminate

Table 3. Descriptive statistics of business outcomes

		All	Average ability			Ability dispersion		
			Low mean<17	Moderate 17≥mean≤21	High mean>21	Low CV<0.15	Moderate 0.15≥CV≤0.30	High CV>0.30
		(1)	(2)	(3)	(4)	(5)	(6)	(7)
Sales (euros)	Mean	902	939	944	762	714	1140	636
	SD	974	909	1111	756	537	1267	478
	Min	0	43	0	42	210	42	0
	Max	5049	3505	5049	2553	2019	5049	1576
Profits (euros)	Mean	24	-18	91	-80	-94	115	-37
	SD	422	347	492	323	391	491	273
	Min	-993	-637	-993	-592	-993	-637	-596
	Max	1619	413	1619	353	508	1619	376
Positive profits (0/1)	Mean	0.57	0.62	0.60	0.45	0.42	0.63	0.62
	SD	0.50	0.51	0.50	0.52	0.51	0.49	0.51
	Min	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	Max	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Profits per share (euros)	Mean	0.62	0.37	1.43	-0.91	-0.85	1.94	-0.46
	SD	7.01	5.71	7.88	6.58	5.31	8.42	5.27
	Min	-13.76	-9.65	-10.13	-13.76	-10.13	-13.76	-8.25
	Max	25.70	9.89	25.70	8.02	8.02	25.70	9.89
N		49	13	25	11	12	24	13

Note: Average and CV of ability reflect at the team level respectively average score and coefficient of variation in scores on Raven's advanced progressive matrices test.

Table 4. Randomization checks at the individual and team level

	CV ability		CV ability^2		Average ability	
A1: Low ability students						
Age	-4.825	(5.894)	16.482	(10.468)	-0.029	(0.068)
Gender	1.591	(1.293)	-4.452*	(2.394)	0.029*	(0.016)
Risk aversion	-3.301	(6.071)	13.287	(12.401)	-0.048	(0.070)
Grade point average	0.227	(0.807)	-1.311	(1.462)	-0.003	(0.010)
A2: High ability students						
Age	7.163	(5.515)	-10.984	(10.448)	-0.057	(0.049)
Gender	-1.682	(1.992)	3.388	(4.653)	-0.020	(0.017)
Risk aversion	-1.639	(6.008)	5.894	(11.588)	0.009	(0.070)
Grade point average	0.755	(0.499)	-0.874	(0.991)	0.005	(0.006)
B: Team level (average)						
Age	-0.048	(5.651)	5.631	(11.485)	-0.009	(0.047)
Gender	-0.537	(2.333)	0.517	(5.366)	-0.009	(0.012)
Risk aversion	-3.195	(6.361)	10.764	(12.887)	-0.009	(0.048)
Grade point average	0.371	(0.593)	-0.789	(1.280)	0.020***	(0.005)
Team size	-1.531	(18.318)	8.077	(39.221)	-0.057	(0.135)

Note: Average and CV of ability reflect at the team level respectively average score and coefficient of variation in scores on Raven's advanced progressive matrices test. In panels A1 and A2 each coefficient comes from a regression at the individual level of the row variable on the column variables, separately for students of low (test score ≤ 18.60) and high (test score > 18.60) cognitive ability (robust standard errors in parentheses). In panel B each coefficient comes from a regression at the team level of the row variable on the column variables (bootstrapped standard errors in parentheses; 1000 replications). ***/**/* denotes significance at the 1%/5%/10%-level.

4 Results

4.1 Main findings

The key prediction of our proposed model in section 2 is that team performance exhibits an inverse U-shaped relationship with ability dispersion. Another prediction of the model is that teams of moderate ability dispersion experience fewer dismissals due to fewer shirking members in those teams. This subsection presents the empirical findings in the order of the propositions of section 2.

Table 5 reports regression results for the effect of ability dispersion on business performance as measured in terms of sales, profits, a binary indicator for positive profits and profits per share (see Proposition 1). In panel A these performance measures are regressed on teams' average test score, their coefficient of variation in test scores and its square (panel B provides results from spline functions). The linear effect of ability dispersion on business performance turns out insignificant in all specifications (not tabulated). Besides standard OLS regression, we employ median and robust (M-estimation) regression to assess whether the results are sensitive to outliers.

---

the design of our field experiment), the analyses in the next section do not include control variables (adding superfluous controls would only reduce the degrees of freedom).

Table 5. Ability dispersion and team performance

	Sales			Profits			Pos. profits	Profits per share		
	OLS (1)	Median (2)	Robust (3)	OLS (4)	Median (5)	Robust (6)	OLS (7)	OLS (8)	Median (9)	Robust (10)
A: Polynomial										
CV ability	14772.5*	4439.5	4422.5	8243.7**	7378.1**	6052.4*	8.952***	106.100*	92.478*	89.344
	(8083.3)	(5549.2)	(5217.3)	(3543.9)	(3200.0)	(3233.6)	(3.163)	(58.466)	(52.612)	(57.936)
CV ability^2	-30305.8*	-9485.4	-9118.5	-16025.3**	-14878.5**	-12441.4**	-16.677**	-207.870*	-196.543*	-178.400
	(17055.2)	(10975.2)	(11357.6)	(6795.2)	(6385.5)	(6276.4)	(6.580)	(112.727)	(104.931)	(115.048)
Avg. ability	25.529	3.449	-2.862	19.528	18.594	4.261	0.004	0.252	0.105	0.115
	(62.982)	(50.458)	(44.774)	(25.975)	(26.966)	(25.490)	(0.032)	(0.461)	(0.475)	(0.467)
Maximum	*0.24*	*0.23*	*0.24*	*0.26*	*0.25*	*0.24*	*0.27*	*0.26*	*0.24*	*0.25*
R^2	*0.09*	*0.04*	*0.03*	*0.15*	*0.08*	*0.11*	*0.14*	*0.09*	*0.08*	*0.07*
B: Spline (0.25)										
1st segment	7377.2*	2629.7	1976.8	3817.7**	3845.9**	2695.7	3.892***	51.036*	46.191	50.538
	(3788.4)	(2486.5)	(2310.0)	(1895.2)	(1583.8)	(1806.3)	(1.508)	(30.705)	(28.296)	(30.742)
2nd segment	-8317.5**	-2818.4	-2247.2	-3352.9**	-4022.5**	-2879.5*	-2.576	-47.116*	-59.334**	-46.582*
	(4060.6)	(3323.5)	(3133.5)	(1694.5)	(1560.1)	(1724.4)	(1.796)	(28.036)	(25.836)	(28.026)
Avg. ability	42.106	29.512	0.560	25.513	36.167	12.321	0.007	0.342	0.209	0.349
	(67.387)	(51.983)	(45.011)	(27.893)	(29.403)	(27.956)	(0.034)	(0.515)	(0.516)	(0.543)
R^2	*0.14*	*0.05*	*0.04*	*0.17*	*0.09*	*0.13*	*0.13*	*0.11*	*0.10*	*0.12*

Note: Based on information from 49 teams. Average and CV of ability reflect at the team level respectively average score and coefficient of variation in scores on Raven's advanced progressive matrices test. OLS, Median and Robust refer to the estimation method. Median and robust specifications for positive profits are excluded since this variable is dichotomous. Bootstrapped standard errors in parentheses (1000 replications). ***/**/* denotes significance at the 1%/5%/10%-level.

Column (1) shows that, given teams' average ability, sales first increase with ability dispersion up to a coefficient of variation in test scores of approximately 0.25 (the sample average is 0.22) and then decrease with ability dispersion. However, columns (2) and (3) indicate that this effect of ability dispersion on sales tends to be inflated by outliers: the point estimates are insignificant when using median and robust (M-estimation) regression estimation techniques. Columns (4) through (6) consistently show an inverse U-shaped pattern for the relationship between ability dispersion and profits. Again, performance is maximized at a coefficient of variation in test scores of about 0.25. The same holds for the probability of profits being positive in column (7), although the degree of ability dispersion where performance peaks marginally increases to a coefficient of variation in test scores of 0.27. The coefficients in columns (8) through (10) corroborate these findings: the effect of ability dispersion on profits per share is described by an inverse U-shape with the optimum at a coefficient of variation in test scores of roughly 0.25. The results from the quadratic specifications in columns (4) through (10) are robust to outliers. Similar results are obtained when we exclude teams' average ability or include higher-order terms for the average ability of teams (not tabulated).

In panel B we estimate spline functions to address asymmetric effects of ability dispersion below (1st segment) and above (2nd segment) a coefficient of variation in test scores of 0.25. The cutoff in our spline functions is obtained by averaging the coefficients of variation in cognitive ability that maximize team performance (according to the quadratic specifications). Results from these spline functions indicate that business performance tends to increase with ability dispersion below a coefficient of variation in test scores of 0.25. If the coefficient of variation in test scores is at least equal to 0.25 all coefficients for the impact of ability dispersion are negative and (with three exceptions) significant. The point estimates in column (5) of panel B imply that raising the coefficient of variation in test scores from 0.20 to 0.25 increases profits by about 200 euros (approximately half of a standard deviation), while profits decrease by roughly the same amount if the coefficient of variation in test scores is further raised from 0.25 to 0.30. In sum, the results of panel B closely resemble the inverse U-shaped pattern from the quadratic specifications of panel A, although the number of teams may slightly limit the precision of its estimates.

Table 6 tests Proposition 2 by estimating the relationship between ability dispersion, dismissals and business performance. The number and incidence of dismissals reflect respectively the number of dismissals per team and whether or not a team has experienced at least one dismissal (dummy = 1 if any). Panel A shows for both number and incidence that teams of moderate ability dispersion are characterized by fewer dismissals and, we infer, less free-riding. Dismissals do not reflect a process whereby teams simply get rid of low ability or high ability students (rather than shirkers) since the relationship between cognitive ability and probability of dismissal at the individual level is insignificant (not tabulated). Moreover, the number and incidence of dismissals are minimized at a coefficient of variation in test scores of approximately 0.24. Note that this minimum almost exactly corresponds with the coefficient of variation in test scores that maximizes business performance (about 0.25). Panel B indicates that fewer dismissals are also positively related to business performance, separately for number and incidence. However, we lack exogenous variation to identify

Table 6. Ability dispersion, dismissals and team performance

A: Dismissals

	Number			Incidence
	OLS	Median	Robust	OLS
CV ability	-14.140***	-18.227**	-13.057**	-8.416***
	(4.154)	(8.399)	(5.523)	(3.260)
CV ability²	29.525***	38.475**	28.681***	17.191***
	(8.136)	(17.091)	(10.264)	(6.489)
Avg. ability	0.018	-0.001	-0.007	0.012
	(0.027)	(0.020)	(0.028)	(0.024)
Minimum	*0.24*	*0.24*	*0.23*	*0.24*
R^2	*0.37*	*0.03*	*0.36*	*0.25*

B: Performance

	Sales			Profits			Pos. profits	Profits per share		
	OLS	Median	Robust	OLS	Median	Robust	OLS	OLS	Median	Robust
Dismissals	-572.6***	-283.5	-278.9*	-193.0**	-176.0	-185.2*	-0.209	-3.270*	-4.392**	-3.120*
(number)	(189.1)	(202.5)	(164.4)	(90.5)	(127.4)	(98.8)	(0.129)	(1.775)	(1.988)	(1.874)
R^2	*0.13*	*0.05*	*0.10*	*0.07*	*0.05*	*0.10*	*0.07*	*0.07*	*0.06*	*0.08*
Dismissals	-774.2***	-355.2	-430.0**	-244.3***	-164.0	-223.4*	-0.259	-4.653**	-4.057	-4.038*
(incidence)	(269.1)	(285.5)	(208.7)	(123.7)	(166.8)	(127.8)	(0.166)	(2.033)	(2.768)	(2.137)
R^2	*0.14*	*0.06*	*0.14*	*0.06*	*0.03*	*0.09*	*0.06*	*0.08*	*0.04*	*0.08*

Note: Based on information from 49 teams. Average and CV of ability reflect at the team level respectively average score and coefficient of variation in scores on Raven's advanced progressive matrices test. All specifications control for team size. OLS, Median and Robust refer to the estimation method. Median and robust specifications for incidence of dismissals and positive profits are excluded since these variables are dichotomous. Bootstrapped standard errors in parentheses (1000 replications). ***/**/* denotes significance at the 1%/5%/10%-level.

a causal impact of dismissals on business performance, since dismissals are obviously endogenous.

To test whether knowledge breadth increases with cognitive ability, we regress the different measures of business performance on a third-order term for ability dispersion (see Proposition 3). Its coefficients turn out insignificant in all specifications (not tabulated), which suggests that knowledge breadth is invariant to cognitive ability in our setting. This is a novel finding, which is corroborated by the (non-tabulated) insignificant point estimates for the average ability of team members. Similar insignificant results are obtained if we regress business performance on average ability and a third-order term for ability dispersion simultaneously (not tabulated).

Consistent with the predictions of our model, the results presented in this subsection show that team performance exhibits an inverse U-shaped relationship with ability dispersion. With average cognitive ability held constant, team performance is maximized at a coefficient of variation in cognitive ability of approximately 0.25 (the sample average is 0.22). Teams of moderate ability dispersion also experience fewer dismissals during the program (i.e., lower degree of shirking members), although this does not chiefly explain why those teams perform better. In contrast to empirical studies which have analyzed settings involving laboratory experiments (Woolley et al., 2010), unskilled work tasks (e.g., Hamilton et al., 2003), or competitive sports (Kahn, 2000), we do not find evidence that average cognitive ability of team members significantly determines performance in a setting that requires coordination on a broad array of complex decision-making tasks entailing the sustained application of their cognitive abilities. Hence, our results suggest that (moderate) dispersion of cognitive ability trumps average cognitive ability in teams comprised of individuals of relatively high cognitive ability that have to complete a complicated business project (see Hong and Page, 2001, 2004). We do not find support for the notion that the lowest-ability member (i.e., the weakest link) or the highest-ability member (i.e., the superstar) in a team significantly affect team performance (Kremer, 1993; Prat, 2002). Results turn out being insignificant too when we relate teams' top 3 or bottom 3 students in terms of cognitive ability to the performance of teams.

4.2 Robustness

Robustness checks in this subsection are conducted by testing other measures of ability dispersion such as teams' standard deviation in ability (Hansen et al., 2006), teams' ratio of the maximum to the minimum ability (Hamilton et al., 2003), and spline functions with three segments of ability dispersion (cutoff levels at a coefficient of variation in test scores of 0.15 and 0.30).

Panels A and B of Table A2 in the appendix also reveal an inverse U-shaped effect of teams' standard deviation in ability and teams' ratio of the maximum to the minimum ability on their performance, although significance levels vary across both measures of ability dispersion.[14] The degree of ability dispersion that maximizes team performance is again slightly above the sample average.

---

[14]The ratio of the maximum to the minimum ability in the team is possibly more sensitive to outliers since this measure of ability dispersion could already be considerably inflated by only one team member of (very) low or high cognitive ability.

Panel C indicates a similar inverse U-shaped pattern for spline functions with teams of low, moderate and high ability dispersion (based on the coefficient of variation in test scores). The point estimates for teams of low ability dispersion (CV<0.15) are positive and relatively large compared to the coefficients for teams of moderate ability dispersion (0.15≥CV≤0.30). For teams of high ability dispersion (CV>0.30) the point estimates are negative and also relatively large in relation to those for teams of moderate ability dispersion. The number of teams, however, limits the precision of these estimates. In sum, none of the robustness checks are at odds with the results previously obtained.

5   Discussion and conclusion

Over the past decades, teams have become increasingly relevant for organizational decision making and performance (Hamilton et al., 2003; Woolley et al., 2010). As a consequence, the composition of teams has become an interesting potential driver of organizational performance. We have studied the effect on organizational performance of a team's composition in terms of cognitive ability, a major determinant of economic behavior and outcomes (e.g., Hanushek and Woessmann, 2008).

To this end, a field experiment was conducted in which teams of undergraduate students start up and manage a real company under identical circumstances. Our experiment is likely to measure the causal effect of ability dispersion on team performance in a setting that closely resembles the functioning of business teams in the longer run (where tasks are diverse and involve complex decision-making).

We propose a model in which greater ability dispersion generates greater knowledge for a team, but also increases the costs of monitoring necessitated by moral hazard. In line with the predictions of our model, team performance exhibits an inverse U-shaped relationship with ability dispersion. Controlling for the average cognitive ability of teams, performance is maximized at a coefficient of variation in cognitive ability of about 0.25 (the sample average is 0.22). Teams of moderate ability dispersion also experience fewer dismissals due to a lower degree of free-riding members, although this does not chiefly explain why those teams perform better.

There is ample opportunity to extend our model, for example, by allowing for richer interactions between team members. That might enable researchers to study other interesting questions such as preferences, beliefs and (re)negotiation in teams. Other limitations relate to the experimental set-up of our study. We exploit exogenous variation in cognitive ability among students who probably lack serious work experience, which may limit the external validity of our findings.

Nevertheless, students in our experiment execute a substantial business project that requires coordination of a broad array of complex decision-making tasks entailing the sustained application of their cognitive abilities. Moreover, students face strong incentives that align their interests with those of the team. All in all, we therefore have grounds to believe that our field experiment is informative about the impact of ability dispersion on the performance of business (management) teams. A next step for future research would be to replicate experiments like this, preferably in real organizations and on a larger scale.

References

Baumol, W. and Strom, R. (2010). *"Useful Knowledge" of Entrepreneurship: Some Implications of the History*, chapter in: D. Landes, J. Mokyr, and W. Baumol (eds.), The Invention of Enterprise: Entrepreneurship from Ancient Mesopotamia to Modern Times, pages 528–541. Princeton: Princeton University Press.

Bors, D. and Stokes, T. (1998). Raven's advanced progressive matrices: Norms for first-year university students and the development of a short form. *Educational and Psychological Measurement*, 58(3):382–398.

Calvin, C., Deary, I., Fenton, C., Roberts, B., Der, G., Leckenby, N., and Batty, G. (2011). Intelligence in youth and all-cause-mortality: Systematic review with meta-analysis. *International Journal of Epidemiology*, 40(3):626–644.

Carpenter, P., Just, M., and Shell, P. (1990). What one intelligence test measures: A theoretical account of the processing in the Raven progressive matrices test. *Psychological Review*, 97(3):404–431.

Cattell, R. (1963). Theory of fluid and crystallized intelligence: A critical experiment. *Journal of Educational Psychology*, 54(1):1–22.

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? *American Economic Review*, 100(3):1238–60.

Dohmen, T., Falk, A., Huffman, D., Sunde, U., Schupp, J., and Wagner, G. (2011). Individual risk attitudes: Measurement, determinants and behavioral consequences. *Journal of the European Economic Association*, 9(3):522–550.

Falk, A. and Ichino, A. (2006). Clean evidence on peer effects. *Journal of Labor Economics*, 24(1):39–57.

Franck, E. and Nüesch, S. (2010). The effect of talent disparity on team productivity in soccer. *Journal of Economic Psychology*, 31(2):218–229.

Gould, E. and Winter, E. (2009). Interactions between workers and the technology of production: Evidence from professional baseball. *Review of Economics and Statistics*, 91(1):188–200.

Grinblatt, M., Keloharju, M., and Linnainmaa, J. (2011). IQ and stock market participation. *Journal of Finance*, 66(6):2121–2164.

Hamel, R. and Schmittmann, V. (2006). The 20-minute version as a predictor of the Raven advanced progressive matrices test. *Educational and Psychological Measurement*, 66(6):1039–1046.

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.

Hamilton, B., Nickerson, J., and Owan, H. (2012). *Diversity and Productivity in Production Teams*, chapter in: A. Bryson (ed.), Advances in the Economic Analyses of Participatory & Labor-Managed Firms, volume 13, pages 99–138. Bingley: Emerald Group Publishing Limited.

Hansen, Z., Owan, H., and Pan, J. (2006). The impact of group diversity on performance and knowledge spillover: An experiment in a college classroom. NBER Working Paper.

Hanushek, E. and Woessmann, L. (2008). The role of cognitive skills in economic development. *Journal of Economic Literature*, 46(3):607–668.

HBO-raad (2010). Feiten en cijfers: Afgestudeerden en uitvallers in het hoger beroepsonderwijs. Technical report, HBO-raad.

Hong, L. and Page, S. (2001). Problem solving by heterogeneous agents. *Journal of Economic Theory*, 97(1):123–163.

Hong, L. and Page, S. (2004). Groups of diverse problem solvers can outperform groups of high-ability problem solvers. *Proceedings of the National Academy of Sciences of the United States of America (PNAS)*, 101(46):16385–16389.

Hoogendoorn, S., Oosterbeek, H., and Van Praag, M. (2013). The impact of gender diversity on the performance of business teams: Evidence from a field experiment. *Management Science*, forthcoming.

Hoogendoorn, S. and Van Praag, M. (2012). Ethnic diversity and team performance: A field experiment. IZA Discussion Paper.

Iranzo, S., Schivardi, F., and Tosetti, E. (2008). Skill dispersion and firm productivity: An analysis with employer-employee matched data. *Journal of Labor Economics*, 26(2):247–285.

Irwing, P. and Lynn, R. (2005). Sex differences in means and variability on the progressive matrices in university students: A meta-analysis. *British Journal of Psychology*, 96(4):505–524.

Kahn, L. (2000). The sports business as a labor market laboratory. *Journal of Economic Perspectives*, 14(3):75–94.

Kandel, E. and Lazear, E. (1992). Peer pressure and partnerships. *Journal of Political Economy*, 100(4):801–817.

Kremer, M. (1993). The O-ring theory of economic development. *Quarterly Journal of Economics*, 108(3):551–575.

Laffont, J. and Martimort, D. (2002). *The Theory of Incentives: The Principal-Agent Model.* Princeton: Princeton University Press.

Lazear, E. (1999). Globalisation and the market for team-mates. *Economic Journal*, 109(454):15–40.

Mas, A. and Moretti, E. (2009). Peers at work. *American Economic Review*, 99(1):112–145.

Mills, C., Ablard, K., and Brody, L. (1993). The Raven's progressive matrices: Its usefulness for identifying gifted/talented students. *Roeper Review*, 15(3):183–186.

Nisbett, R., Aronson, J., Blair, C., Dickens, W., Flynn, J., Halpern, D., and Turkheimer, E. (2012). Intelligence: New findings and theoretical developments. *American Psychologist*, 67(2):1–30.

Oosterbeek, H., Van Praag, M., and IJsselstein, A. (2010). The impact of entrepreneurship education on entrepreneurship skills and motivation. *European Economic Review*, 54(3):442–454.

Papps, K., Bryson, A., and Gomez, R. (2011). Heterogeneous worker ability and team-based production: Evidence from major league baseball, 1920–2009. *Labour Economics*, 18(3):310–319.

Prat, A. (2002). Should a team be homogeneous? *European Economic Review*, 46(7):1187–1207.

Raven, J. (2000). The Raven's progressive matrices: Change and stability over culture and time. *Cognitive Psychology*, 41(1):1–48.

Raven, J., Raven, J., and Court, J. (1993). *Manual for Raven's Progressive Matrices and Vocabulary Scales: Section 1 General overview.* Oxford: Psychologists Press.

Raven, J., Raven, J., and Court, J. (1998). *Manual for Raven's Progressive Matrices and Vocabulary Scales: Section 4 Advanced Progressive Matrices.* Oxford: Oxford Psychologists Press.

Rosen, S. (1981). The economics of superstars. *American Economic Review*, 71(5):845–858.

Rushton, J. and Jensen, A. (2005). Thirty years of research on race differences in cognitive ability. *Psychology, Public Policy and Law*, 11(2):235–294.

Spearman, C. (1927). *The Abilities of Man.* New York: Macmillan.

Stewart, G. (2006). A meta-analytic review of relationships between team design features and team performance. *Journal of Management*, 32(1):29–55.

Woolley, A., Chabris, C., Pentland, A., Hashmi, N., and Malone, T. (2010). Evidence for a collective intelligence factor in the performance of human groups. *Science*, 330(6004):686–688.

Appendix Table A1. Team characteristics

#	Name	Average ability	CV ability	Team size	Sales (euros)	Profits (euros)	Profits/ share	Description of product/service
1	Achter de rits	17.64	0.33	11	533.95	78.20	1.56	Sustainable fair trade boxer short
2	Be Portable	16.00	0.22	10	1179.58	288.30	4.65	Portable durable water bottle
3	Berry Emotions	21.45	0.26	12	88.24	-591.58	-13.76	Personalized telephone cover
4	Blue Empire	14.09	0.19	13	418.07	-75.44	-1.80	Fashionable bluetooth bracelet
5	Cleaning Env. Friendly	14.75	0.31	8	415.76	3.72	0.06	Long-lasting durable paper towel
6	Chori-Chori	16.62	0.20	15	588.24	-441.00	-3.94	Trendy customized Indian jewelry
7	Comb. Young E'ship	15.13	0.19	9	894.12	119.90	2.45	Whistling operated key finder
8	Covers 'n Cases	19.36	0.37	16	338.87	56.10	1.52	Covers and cases for telephones
9	Easylife	19.13	0.32	12	207.18	-288.79	-8.25	Automated toothpaste dispenser
10	Flashing Productions	16.70	0.31	12	615.88	90.34	1.96	Fluorescent nightlife bracelet
11	FresHES	15.08	0.20	15	3505.35	145.51	3.10	Wine bottle shaped umbrella
12	Freshual	22.08	0.15	13	1599.66	352.81	4.58	Convenient compact key torch
13	Green	19.50	0.21	10	524.37	111.81	4.66	Unbreakable vase for flowers
14	Green Solutions	14.07	0.31	14	1576.38	375.65	9.89	Environmental awareness products
15	Horanje	17.38	0.33	14	463.87	36.55	0.52	Gift bag with Dutch souvenirs
16	Ibrella	17.18	0.18	12	238.66	-473.79	-6.97	Easy-to-wear mini umbrella
17	I-corporation	16.27	0.22	15	43.15	-636.81	-9.65	Convenient eye-drop dispenser
18	I'Magine	18.83	0.21	15	3267.66	702.89	5.81	Fashionable Brazilian flip-flops
19	Innovart	22.40	0.16	12	42.02	-8.40	-0.20	Broker/intermediary in art
20	Jikoni	18.56	0.20	10	861.05	54.26	1.36	Innovative kitchen appliances
21	Kamikare	17.50	0.18	10	577.98	27.05	0.52	Easy-to-use hair styling product
22	K'deau	19.83	0.32	17	878.00	120.00	1.74	Gift box with various gadgets
23	Keyen	19.00	0.13	13	1525.79	256.11	3.41	High-quality key finder device
24	Lightdancer	16.17	0.47	13	139.79	-595.77	-7.84	Jewelry with built-in laser beam
25	MagicTouch	21.86	0.08	7	539.50	-532.96	-5.49	Compact women-friendly vibrator

Appendix Table A1. (continued)

#	Name	Average ability	CV ability	Team size	Sales (euros)	Profits (euros)	Profits/ share	Description of product/service
26	Muffle	17.78	0.36	11	1320.76	331.09	6.13	Trendy coffee to-go cup sleeve
27	Nailtastic	17.78	0.15	10	284.87	-284.63	-7.49	Nailpolish applicator help kit
28	New Young Pros	20.43	0.30	10	1476.48	794.00	12.41	Anti-slip mat for multiple devices
29	Razor2Go	18.42	0.08	13	387.68	-113.31	-1.83	Comfortable facial hair remover
30	ReWrite	14.75	0.25	11	1600.45	413.39	6.36	T-shirt with rewritable tekst area
31	Rising Sun	14.78	0.30	11	577.71	-113.11	-4.71	Mini high-definition pen camera
32	Sleevies	20.13	0.25	10	5048.50	1619.39	25.70	Public transport chip card holder
33	Stud. Offer Products	22.22	0.18	9	144.18	-217.74	-4.35	Aids awareness-raising key chain
34	Sunny & Co	14.00	0.25	10	653.19	194.73	4.23	Wristband with sunscreen notifier
35	SunVision	20.40	0.17	10	391.66	-197.47	-5.81	Affordable sunglasses on strength
36	Swops	23.22	0.12	9	557.15	112.69	1.28	Trendy digital slim sports watch
37	Cardholder Company	18.11	0.12	9	252.10	-140.00	-5.60	Cardholder wallet with mirror
38	The Mug	20.70	0.15	13	2019.27	508.46	5.98	Comfortable coffee to-go mug
39	TieSupply	20.38	0.07	10	433.74	-242.86	-1.80	Tie roll cylinder in a travel case
40	Toccare	21.13	0.12	15	463.85	-225.90	-5.65	Key chain with digital business card
41	Winterwarmth	20.11	0.21	13	893.28	480.50	14.56	Multi-purpose hot and cold gel pad
42	Wristbands4World	17.69	0.14	14	596.00	50.88	1.64	Environmental awareness wristband
43	Xihe	20.70	0.09	11	860.50	-993.02	-10.13	Innovative solar battery charger
44	XXL	21.08	0.19	12	2552.91	238.86	2.91	HES University hooded sweatshirt
45	BubbleBeat	17.75	0.35	8	0.00	-225.00	-3.88	n/a
46	Cotton&Class	21.25	0.30	10	1201.00	-352.00	-4.69	n/a
47	Decent	19.25	0.08	13	210.00	-2.00	-0.08	n/a
48	InOffice	21.11	0.15	10	475.00	154.00	7.33	n/a
49	WatchUs	21.75	0.10	13	716.67	192.50	8.02	n/a

Note. Average and CV of ability reflect at the team level respectively average score and coefficient of variation in scores on Raven's advanced progressive matrices test (excluding students whose test score is unknown). The number of students whose test score is unknown amounts to 60 (out of 573 students). Team size reflects the size of teams at baseline. Missing or incomplete descriptions of a team's product or service are indicated by 'n/a' (not available).
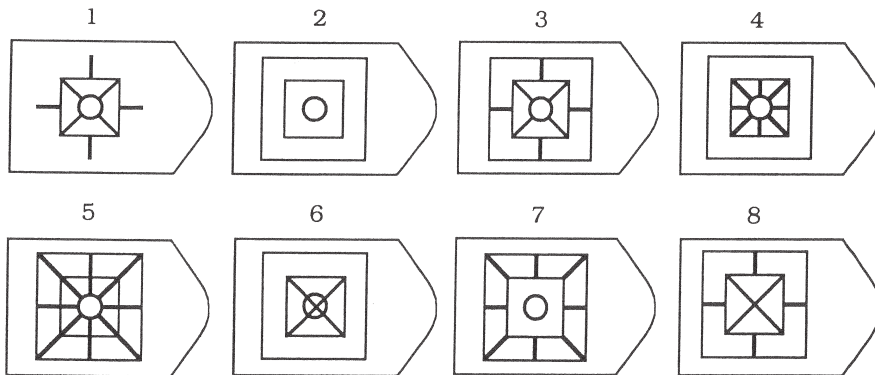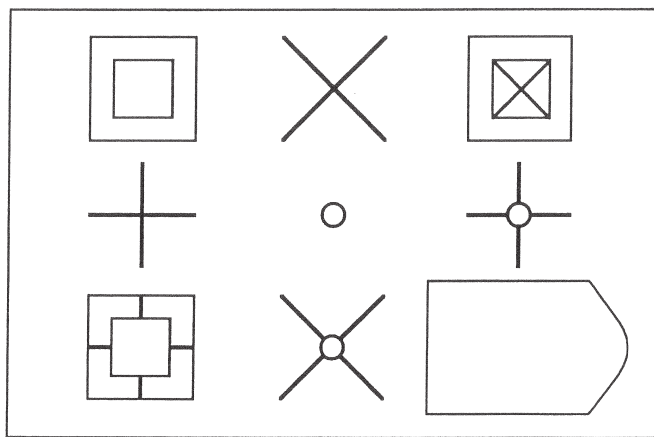
Appendix Figure A1. Example of a figure from Raven's advanced progressive matrices test

Appendix Table A2. Other measures of ability dispersion and team performance

	Sales			Profits			Pos. profits	Profits per share		
	OLS (1)	Median (2)	Robust (3)	OLS (4)	Median (5)	Robust (6)	OLS (7)	OLS (8)	Median (9)	Robust (10)
A: Standard deviation										
SD ability	955.5**	370.4	239.7	548.5***	538.0**	412.6**	0.641***	7.529**	6.324*	6.010**
	(446.4)	(286.2)	(272.0)	(202.6)	(206.4)	(169.9)	(0.195)	(3.131)	(3.650)	(2.765)
SD ability²	-107.6**	-43.7	-27.7	-59.4***	-65.1**	-46.9**	-0.068***	-0.827**	-0.801*	-0.679**
	(48.4)	(33.4)	(32.3)	(22.0)	(24.6)	(19.2)	(0.024)	(0.339)	(0.424)	(0.314)
Avg. ability	18.841	30.451	-6.546	11.389	-2.816	-0.248	-0.008	0.163	-0.083	0.013
	(57.584)	(44.307)	(36.635)	(20.706)	(23.714)	(18.422)	(0.027)	(0.386)	(0.462)	(0.364)
Max./Avg.	*4.44/2.17*	*4.24/3.91*	*4.33/3.91*	*4.62/3.91*	*4.13/3.91*	*4.40/3.91*	*4.71/3.91*	*4.55/3.91*	*3.95/3.91*	*4.43/3.91*
R²	*0.08*	*0.03*	*0.02*	*0.15*	*0.06*	*0.12*	*0.16*	*0.10*	*0.07*	*0.08*
B: Ratio maximum/minimum										
Ratio ability	2208.6**	1037.6	709.8	843.1	386.5	512.0	0.849	10.392	7.459	7.608
	(1043.0)	(787.2)	(788.1)	(550.5)	(571.2)	(606.1)	(0.683)	(9.255)	(10.154)	(10.879)
Ratio ability²	-403.5*	-181.8	-131.9	-151.5	-58.3	-91.0	-0.140	-1.851	-1.127	-1.342
	(212.2)	(164.0)	(167.3)	(114.9)	(124.1)	(130.6)	(0.146)	(1.908)	(2.214)	(2.352)
Avg. ability	43.312	40.983	7.209	18.076	13.565	3.956	0.003	0.242	0.389	0.089
	(56.444)	(41.184)	(41.247)	(23.082)	(26.870)	(22.130)	(0.033)	(0.438)	(0.530)	(0.436)
Max./Avg.	*2.74/2.17*	*2.85/3.91*	*2.69/2.17*	*2.78/2.17*	*3.31/2.17*	*2.81/2.17*	*3.03/2.17*	*2.81/2.17*	*3.31/2.17*	*2.83/2.17*
R²	*0.10*	*0.05*	*0.04*	*0.08*	*0.03*	*0.04*	*0.08*	*0.04*	*0.03*	*0.03*
C: Spline (0.15 and 0.30)										
1st segment	9022.6*	2193.0	186.9	5166.3*	5085.9	3918.4	7.297**	55.414	60.502	51.447
(N=12)	(4649.3)	(4643.2)	(4629.7)	(2702.6)	(3112.3)	(2754.7)	(2.960)	(40.254)	(53.613)	(42.579)
2nd segment	1045.1	408.3	1805.1	1120.6	481.3	441.3	0.759	17.213	1.709	10.826
(N=24)	(3232.4)	(2600.9)	(2350.0)	(1493.7)	(1521.2)	(1687.2)	(1.814)	(27.227)	(32.648)	(31.896)
3rd segment	-8014.1	-2887.7	-4484.7	-4162.1	-4548.3	-3429.2	-2.799	-57.507	-61.100	-48.734
(N=13)	(9125.7)	(7342.2)	(6849.5)	(3329.1)	(4084.8)	(4576.6)	(4.542)	(65.005)	(71.815)	(109.826)
Avg. ability	21.255	8.269	-3.939	17.804	10.606	3.567	0.001	0.227	0.130	0.085
	(65.578)	(49.193)	(42.950)	(26.909)	(29.171)	(28.291)	(0.033)	(0.477)	(0.550)	(0.511)
R²	*0.08*	*0.04*	*0.06*	*0.14*	*0.07*	*0.10*	*0.14*	*0.08*	*0.08*	*0.07*

Note: Based on information from 49 teams. In panel A average and SD of ability reflect at the team level respectively average score and standard deviation in scores on Raven's advanced progressive matrices test. In panel B average and ratio of ability reflect at the team level respectively average score and ratio of the maximum to the minimum score on Raven's advanced progressive matrices test. In panel C average and CV of ability reflect at the team level respectively average score and coefficient of variation in scores on Raven's advanced progressive matrices test (with cutoffs at a coefficient of variation in test scores of 0.15 and 0.30). OLS, Median and Robust refer to the estimation method. Median and robust specifications for positive profits are excluded since this variable is dichotomous. Bootstrapped standard errors in parentheses (1000 replications). ***/**/* denotes significance at the 1%/5%/10%-level.