Title: Identifying management skill

Appendix: Detailed Pre Analysis Plan

Not for public presentation

Contents

1. Background and research questions

Managers are central to economic production. In the US, managers make up 18% of the workforce and receive nearly 30% of total compensation (US Bureau of Labor Statistics). In the field, firm productivity is affected by the quality of managers' *allocative skill* (Minnie, 2023) and *social skill* (see, for example, Kim 2020). In this paper, we conceptualize *management skill* as a combination of allocative and social skill.

Identifying management skill

How can firms identify good managers? Existing methods are imperfect. One strategy involves observing manager performance in the field (e.g. Bertrand and Schoar, 2003). While this direct approach has high ecological validity, it is difficult to identify the causal contribution that managers make, as managers are not randomly assigned to teams or firms. Moreover, firms only observe the management performance of people who are promoted to manager. Many great managers – especially from non-traditional backgrounds – may go unnoticed.

Another approach is to use self-reported measures of managerial skill (e.g. Beenan 2019). These can be effective for academic research but are easily gamed in high-stakes assessments. In response, firms may use basic assessments of cognitive ability which are difficult to game – e.g. tests of spatial or analogical reasoning. However, these broad instruments typically have limited explanatory power to predict managerial performance.

Finally, firms may assess management potential by using the judgment of existing managers or workers. This provides useful information but there is often a slow feedback loop between promotion decisions and outcomes which makes it difficult to train intuition (Kahnemann and Klein, 2009). In many cases, this feedback loop isn't closed as management performance is not systematically measured.

Why would better methods to identify management skill be helpful?

Limitations in measuring management skill may hamper firms' ability to hire and develop effective managers. Given the growing importance of management, this may have substantial implications for productivity and growth.

A failure to identify management skill may also exacerbate inequality – particularly gender inequality. In the absence of skill measures, other mechanisms determine who becomes managers. One such mechanism is self-nomination, which typically favours men. For example, Chakraborty & Serra (2022) find women are less likely to self-select into leadership positions when team members can talk back to managers and there is a possibility of backlash. These studies imply that with men being more likely to self-select into leadership than women, the process through which managers are allocated to teams can perpetuate the existing gender leadership gap. People who self-nominate for management roles may not make good leaders. Brunell et al. (2008) and Nevicka et al. (2011) find that narcissists tend to emerge as managers in groups and are often judged positively by their team members, even when they negatively impact team performance (Nevicka et al. 2011).

Research questions

Our experiment aims to address two sets of questions.

RQ1: Can we identify good managers and if so, what skills do they have?

a. Can some managers reliably improve team performance (conditional on task-specific skills of team members)?
b. How important are management skills compared to task-specific skills?
    - How much do groups improve when they are randomly assigned someone with strong management skills? How does this compare to a situation where groups are randomly assigned a manager with strong task skills?
c. What characteristics and skills are associated with good managers?
d. Do good managers have different skills from good workers?

RQ2: How should managers be selected?

a. Do people who nominate themselves for management roles have above-average management skills?
b. What is the relationship between gender, management skills, and self-nomination?
c. How do the managerial skills of the top-third (and bottom-third) of people who nominate themselves compare with the managerial skills of the top-third of managers that emerge from randomized assignment (in other words, how do the 2 distributions compare)?
d. Do managers have a good understanding of the skills required to be a good manager? Do workers?

2. Experimental Design and Measures

2.1 Design

The experiment has two main parts:

1. Individual testing. This will take place both at home and at the lab. Individual tests at the lab will occur immediately before and immediately after group testing.
2. Group testing, conducted at the Essex lab.

We discuss these two elements in turn.

2.1.1. Individual testing

Individual tests will focus on **three areas**:

1. Task-specific skills: how well do individual solve problems on their own?
    a. We administer three short tests assessing: numeracy, analytical reasoning, and spatial reasoning
2. Potential predictors of manager and teamwork skill
    a. Allocative skill
        i. Assignment Game
        ii. Indecisiveness Inventory (II)

    b. Measures of social skill
        i. Emotional perceptiveness (RMET)
        ii. Political Skill Inventory (PSI)
3. Broad measures of individual characteristics
    a. Demographic questionnaire
    b. Big5 inventory
    c. IQ (via Culture Fair Intelligence Test)

These individual tests will take place at **three different points in time**:

a. Tests to be taken at home, before the group session
    - Assignment Game
    - RMET
    - IQ (CFIT)
b. Tests to be taken at the lab, immediately prior to group tasks
    - Allocation-task skills (numerical/spatial/analytical reasoning)
    - NB: at this point we will also elicit people's preferences for being a manager
c. Tests to be taken at the lab, immediately after group tasks
    - Demographic Questionnaire (age/ethnicity/education)
    - PSI
    - II
    - Big5

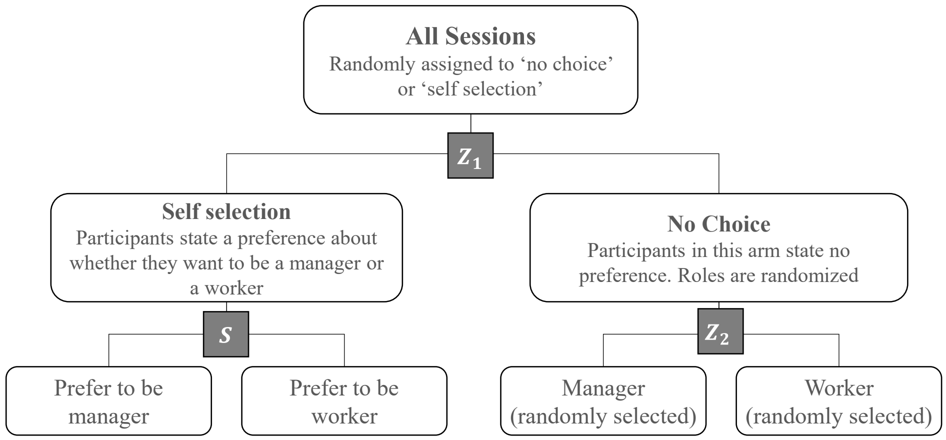
These tests are described in detail below.

2.1.2 Group testing

Participants will be invited to sign up for 1 lab session at the University of Essex lab. Each session will last roughly 90-120 minutes. A majority of this time will be spent on group testing. Participants will work in groups of 3 on an allocation task (described in detail below).

Prior to the start of each lab session, we randomise the session to either the 'self selection' or 'no choice' treatment. Participants in the self selection condition rate on a scale 1/10 their preference for being a manager in the remainder of the experiment. Those with the strongest preference are assigned to be the manager (if there is a tie, roles will be randomised) with all other participants being assigned to the role of 'worker'. Those in the 'no choice' condition will also be asked for their preference to be assigned to the role of manager but ultimately their role will be randomized. This process is summarized in Figure *1*.

Figure 1 – Assigning participants to roles



1. Z_1 is the random variable that determines whether a session will be 'no choice' or 'self selection'. Sessions are paired, and within each pair of sessions we randomly select one to be a 'self selection' session and the other to be a 'no choice' session.
2. Z_2 is the random variable that determines roles in the 'no choice' sessions. We randomly select $\frac{n}{3}$ participants in a session to be a manager, where n is the number of participants in the session (minimum 12, maximum 15).
3. S is the mechanism that determines roles in the self-selection condition. For participant i, S_i is equal to 1 if that participant's 'preference to be a manager' is greater than n/3 of the other participants in the session. If preferences are equal the tie is broken by random allocation.

In the self-selection condition will be presented with the following information to help them select a role:

Figure 2 – Instructions to participants who are assigned to the "self selection" condition

Instructions Group Tasks

In this round of the experiment you will either be a "Manager" or a team member.

Do you want to be a manager?

Managers are responsible for directing the group and making final decisions. The role will involve communicating with your team members, delegating, collating information and making decisions

Manager's whose groups perform in the top 40% of all groups will earn substantially more than other team members in the session. Manager's whose group performance is in the bottom 40% will have substantially lower earnings than other team members in the session

On average, Managers and team members will receive the same payment.

You should choose the role that best fits your skills.

Indicate how much you'd like to be manager in the upcoming experiment, on a scale of 1 (I really don't want to be a manager) to 10 (I really do want to be the manager). Please be as honest as possible, as your preference could be taken into account. However, also note that we cannot guarantee you will get your preferred role

On a scale of 1 to 10, where 1 indicates a strong preference for NOT being the Manager and 10 indicates a strong preference for being the Manager, how much do you want to be the Manager?

1 ○○○○○○○○○○ 10

Next

In the 'no choice' condition participants are shown an identical screen so we are able to elicit preferences to be a manager in both treatments. To avoid deception, participants are told that their preferences could be taken into account and that we cannot guarantee their preferred role.

As noted in the materials, participants in the manager role have stronger performance incentives than workers. This is made known to the participants. Managers whose groups perform in the top 40% of all groups will earn more than workers (£12, compared to £8 for workers). Managers whose group performance is in the bottom 40% of groups will have lower earnings than workers (£4).

During a session, each participant will be randomly assigned to 4 groups, each containing three people. Each group will complete a single Group Assignment task.

2.1.4 Randomization

Randomization takes place within each session. There are two forms of randomization, randomization of sessions to treatment or control and randomization of groups. We explain each in turn:

Randomization of Treatments/Control
We randomly allocate each session to either be a 'self selection' or 'no choice' treatment. Sessions will be equally split across these two conditions.

Randomization of Groups
To run a study 'session' there must be at least n=12 participants. Ideally, there will be n=15 participants at a time (this is our aim). Workers are randomly assigned to managers, with the following constraints: if possible, we do not consider groupings in which workers are allocated to

a manager they have worked with before. Further, workers are allocated to the same manager at most once. To achieve this (when n=15), each manager is randomly given a number from 1:5, and each worker is randomly given a number from 6:15. Group allocations then proceed as follows:

	Round 1		Round 2		Round 3		Round 4	
Manager	W1	W2	W1	W2	W1	W2	W1	W2
1	6	15	7	13	8	11	9	14
2	7	14	8	12	9	15	10	13
3	8	13	9	11	10	14	6	12
4	9	12	10	15	6	13	7	11
5	10	11	6	14	7	12	8	15

W1 and W2 are workers in each group. Note, this assumes a session of 15 participants.

Assuming 15 people in a session, workers will be allocated to a manager at most once.

2.1.5 Sample

The sample will come from the Essex Lab sample pool, with the condition that each subject can only participate once. In addition, the sample will be comprised of students under the age of 35.

2.2 Individual measures

This section describes our measures in detail.

2.2.1 Task-specific skills: how well do individuals solve problems on their own?

Allocation task skills

The allocation task is made up of three modules: numerical; spatial and analogical reasoning. These tasks were selected as existing research suggests there is a low cross correlation between tasks (e.g., Chabris, 2007, Haier et al. 2009), which increases the scope to make efficient allocation decisions based on comparative advantage. We now briefly describe each of the three modules.

Numerical

To assess mathematical reasoning, we use the number sequence task. It is a common cognitive assessment tool used to measure a person's ability to understand and manipulate numbers in a logical sequence. The task typically consists of a series of number sequences, with each sequence having a missing number. The subject is asked to fill in the missing number based on the logic or pattern of the sequence. For example:

4 | 8 | 12 | 16 | ?

This task is used to assess a wide range of cognitive abilities, including mathematical reasoning, problem-solving, and attention to detail. It is often used as a measure of mathematical ability and potential for success in fields such as engineering, finance, and science as part of aptitude tests, psychometric evaluations, and IQ assessments during the job application process. For instance, they are used in the Wonderlic battery of tests, Johnson O'Connor battery, Criteria Cognitive Aptitude Test and the SHL Verify Ability Tests.

Spatial

A spatial reasoning task is a type of cognitive task that assesses an individual's ability to manipulate and visualize objects in two or three dimensions. It involves understanding and working with spatial relationships, such as size, shape, orientation, and position. Examples of spatial reasoning tasks include mental rotation tasks, where a person is asked to imagine rotating a 3D object in their mind, and paper folding tasks, where a person is shown an image of a piece of paper with a folded crease and asked to visualize how it would look if it were unfolded. For example:

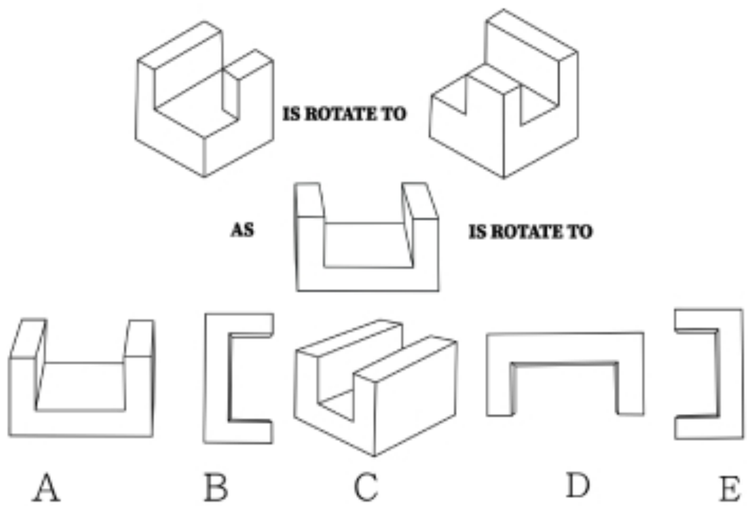


Figure 3: Example of spatial question

Spatial reasoning tasks are commonly used in job application assessments to evaluate a person's ability to understand and manipulate spatial information. These tasks are often used to assess a candidate's potential for success in fields such as engineering, architecture, graphic design, and other technical jobs. Questions similar to this are used across multiple job assessment tasks including Johnson O'Connor, Harver cognitive ability test, Wonderlic battery of tests and the University Clinical Aptitude Test.

Analytical

The analytical reasoning task is a set of tests that assess an individual's ability to understand and analyze language problems and solve analogies. For example:

Cat is to Meow, as Dog is to:

A. Tail

B. Wag
C. Woof
D. Puppy
E. Walk

Analytical reasoning tasks are commonly used in job assessments, particularly for positions that require strong language and communication skills.

Scoring the individual skill tasks

Subjects are given 4 minutes to solve each module. They receive 1 point for each correct answer and their score reduces by 0.5 for each incorrect answer. Participants are made aware that they lose points for incorrect answers.

2.2.2 Potential predictors of manager skills / team membership skills

We measure several potential predictors of management and teamwork skills. These are broken down into two buckets. In each case we include a skill measure, and a self-reported measure.

1. Measures of allocative skills
    a. Assignment Game (AG)
    b. Indecision Inventory (II)
2. Measure of social skill
    a. Reading the Mind in the Eyes (RMET)
    b. Political Skill Inventory

Measures of allocative skill

<u>*The Assignment Game (AG)*</u>

In the Assignment Game (Caplin et al., 2023), participants play the role of a manager who must assign a list of tasks to a group of fictional workers. In each assignment problem there are 3 or 4 tasks and a matching number of workers. Each worker must be assigned to only 1 task, and all tasks must be assigned. The goal is to assign workers so that the output of the team is maximized. Each problem has a single unique solution in which workers are optimally allocated to tasks. To succeed at the task, participants must:

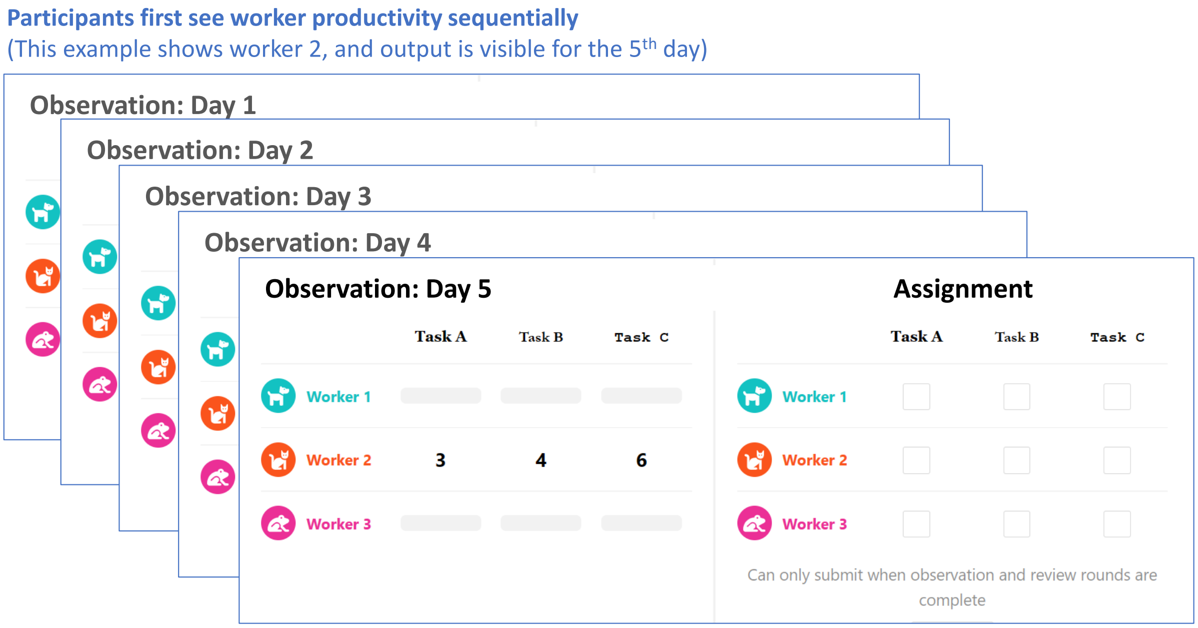
- understand comparative advantage (at least in an intuitive sense) and make use of this idea in concrete allocation decisions
- avoid biases that undermine numerical decision making (e.g. anchoring)
- deal successfully with a highly complex information environment

Observing workers' skills

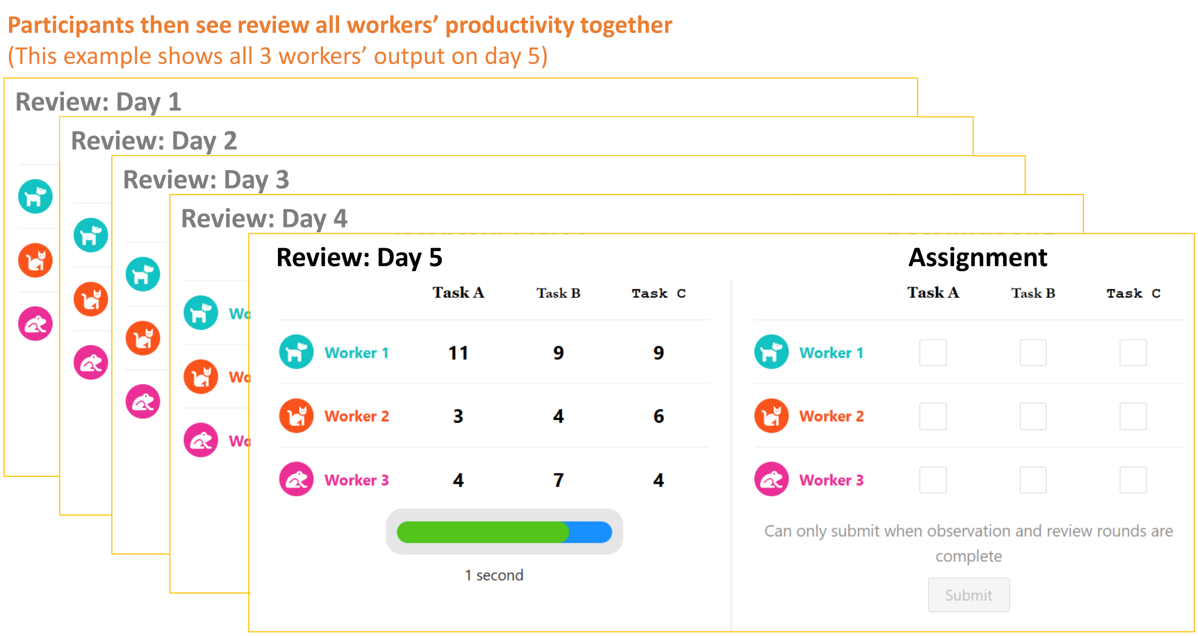
In the first part of the assignment game, participants observe the output of fictional workers. Each worker has a productivity schedule over the tasks and participants are shown multiple draws of workers' output for each task. Participants are told that "workers have good days and bad days" and that as manager their job is to figure out "how good workers are at different tasks ON AVERAGE".

The observation period presents information about worker productivity in two ways. Participants are initially shown multiple draws of each worker's output *by worker* (the outputs worker 1, then worker 2 etc; see top panel of Figure 1). A worker's output for all tasks on a given day is displayed for 1-2 seconds. Next, there is a review period in which information about all workers is presented simultaneously (see bottom panel of Figure 4; note that the review repeats information that participants have already seen). Each review table is presented for 3-4 seconds.

Figure 4 - Observation period in the Assignment Game



[Figure 4 continued over the page]



Assigning tasks to workers

Participants need to assign all the tasks so that each worker has 1 task. The rules are as follows:

- Participants can assign workers at any point during the game, including the observation period (see Figure 5, which shows a participant making an initial assignment for all three workers on day 4 of the observation period for worker 3)
- Participants can change their assignments at any time
- After the observation period ends, participants have 10 seconds to finalize their assignments. During this period, participants lose access to worker productivity information (see Figure 6)

Figure 5 - Participants can assign workers at any time (incl. in the observation period)

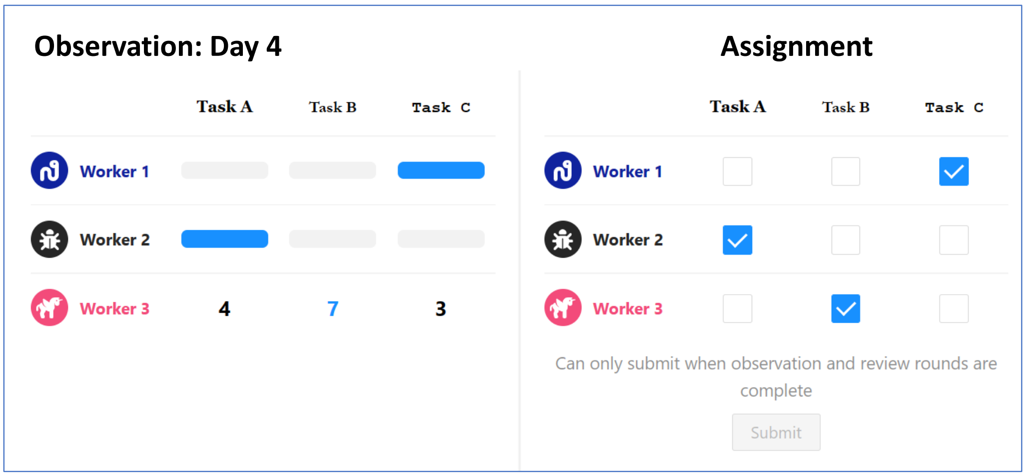
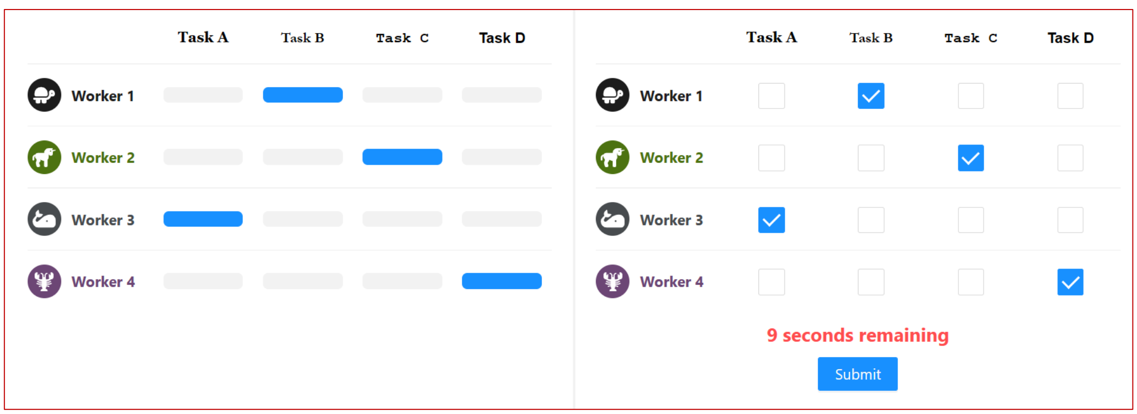


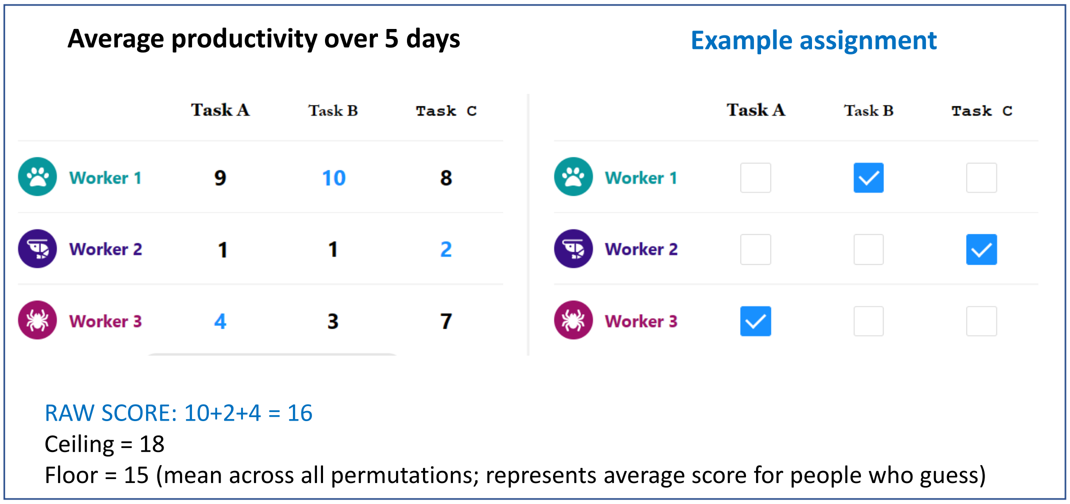
Figure 6 - Final submission period in Assignment Game



Scoring the assignment game

Scores are based on the average productivity of workers. For example, suppose the left panel of Figure 7 gives the **average** worker productivity across the 5-day observation period. If a participant chooses the assignment on the right panel of the figure (Task A to worker 3, Task B to worker 1, and Task C to worker 2) the raw score would be 16 = 4+10+2. Participant scores can be compared to two thresholds: a ceiling score (i.e. the optimal assignment) and a floor (calculated as the expected score from random guessing).

Figure 7 - Example of how scoring works in the Assignment Game



Indecisiveness inventory (II)

The II is a widely used scale to assess how individuals approach decision making. Specifically, it captures if the individual has a tendency to postpone making decisions. In the survey module, participants report the extent to which they agree with statements on various decision making styles and tendencies, "I try to put off making decisions" or "I always know exactly what I want". We will administer the 15-item form outlined below (Frost & Shows 1993).

How strongly do you agree or disagree with the following statements? Please indicate your answer on a scale from 1 to 5. A 1 means "Disagree strongly" and a 5 means "Agree strongly" You can use any number between 1 and 5 to indicate where you fall on the scale, using

1 = Strongly disagree 2 = Disagree 3 = Neither agree nor disagree 4 = Agree 5 = Strongly agree

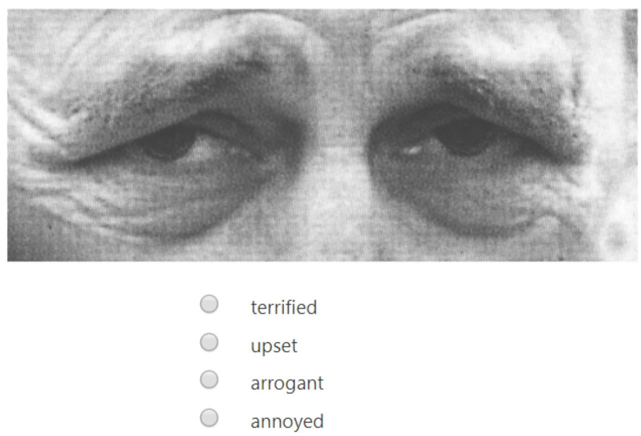
1	I try to put off making decisions
2	I always know exactly what I want
3	I find it easy to make decisions
4	I have a hard time planning my free time
5	I like to be in a position to make decisions
6	Once I make a decision, I feel fairly confident that it is a good one
7	When ordering from a menu, I usually find it difficult to decide what to get
8	I usually make decisions quickly
9	Once I make a decision, I stop worrying about it
10	I become anxious when making a decision
11	I often worry about making the wrong decision
12	After I have chosen or decided something, I often believe I've made the wrong choice or decision
13	I do not get assignments done on time because I cannot decide what to do first
14	I have trouble completing assignments because I can't prioritise what is most important

15	It seems that deciding on the most trivial thing takes me a long time

Measures of social skill

Reading the Mind in the Eyes (RMET)

RMET tests participants' ability to recognize emotions in others and, more broadly, their 'theory of mind' (i.e. their ability to reason about the mental state of others, Baron-Cohen et al., 2001). The test presents participants with photos of faces, cropped so that only the eyes are visible (see example below). For each set of eyes, participants are asked to choose which emotion, from four options, best describes the person in the image. Definitions of all the words are made available via links to an online dictionary.[1]



Relative to other measures of emotional perceptiveness, RMET has good psychometric properties and can be administered quickly (Pinkham et al., 2013). Results from Weidmann and Deming (2020) suggest that RMET is a good predictor of whether or not someone is a Team Player (i.e. someone who makes a positive contribution to group problem solving, conditional on their task-specific skill). Similar results from Woolley et al. (2010) and Engel et al. (2014) suggest that a group's endowment of RMET is perhaps the strongest known predictor of collective intelligence. Individual skill in emotional perception appears to be associated with improvements in group problem solving. In this study, we will examine whether it is specifically associated with management skill and teamwork skill.

Political Skill Inventory (PSI)

The PSI uses self-reports to measure 'political skill' defined as "the ability to effectively understand others at work, and to use such knowledge to influence others to act in ways that enhance one's personal and/or organizational objectives" (Ahearn et al., 2004). This quality is often described as "savvy" (Ferris et al., 2005). In validation work, the PSI was shown to be unrelated to general mental ability, but predictive of managers' performance ratings (Ferris et al., 2005). Managers' scores on the PSI were also found to be a predictor of team performance in a

[1] We use a slightly shortened version of the test. Rather than 36 items, we use 26. We removed 5 female and 5 male images. This was primarily done to save time in what was becoming an onerous battery of tests.

field study (Ahearn et al., 2004). We argue that the PSI is better at capturing social skills that are relevant for interactions with strangers in a lab.

Participants will be asked the items from the shortened PSI described in Ferris et al. (2005).[2]

Item	Prompt
1	I am able to make most people feel comfortable and at ease around me.
2	I am able to communicate easily and effectively with others
3	It is easy for me to develop good rapport with most people.
4	I understand people very well
5	I am particularly good at sensing the motivations and hidden agendas of others.
6	When communicating with others, I try to be genuine in what I say and do.
7	I am good at getting people to like me.
8	It is important that people believe I am sincere in what I say and do.
9	I try to show a genuine interest in other people.
10	I have good intuition or savvy about how to present myself to others.
11	I always seem to instinctively know the right things to say or do to influence others.
12	I pay close attention to people's facial expressions.

2.2.3 Broad measures of individual characteristics

We measure a number of other potential important individual characteristics at the end of the individual stage of the experiment. We explain these below

Background/demographic questionnaire

We will ask the following questions:

1. Age: ___________________________

2. Gender: 0. Female 1. Male. 2 Prefer not to say

3. How many years of work experience do you have?

4. Which program are you enrolled in?
   i. Undergraduate
   ii. Graduate

5. What ethnicity do you most identify with?
   i. Asian or Asian British
   ii. Black, Black British, Caribbean or African
   iii. Mixed or multiple ethnic groups
   iv. White

[2] The items that relate to the "networking" subscale are removed, as these are not relevant to our lab setting

v. Other ethnic group
vi. Prefer not to say

6. How willing are you to take risks, in general?
[Options range between 0 and 10]

This questionnaire will also include a question about participants' perceptions of important skills for 'managers' to have to succeed in our experiment. This will be asked after the conclusion of the group session.

We also ask individuals a set of questions about the experiment they just completed.
In particular, we ask them to imagine they were in charge of **selecting managers** for the task they just completed. We then ask them "Which of these skills would be most important to you, to find a great manager? (Put a '1' next to the skill you think is most important; a '2' next to the skill you think is second in order of importance and so on. The ordering of the set are randomized.)

a. IQ (being generally good at solving problems)
b. Task skills (being specifically good at numerical, spatial and analogical puzzles)
c. Social skills (being good at interacting with people)
d. Allocative skills (being good at making decisions about who should do what)"

Big5 inventory
Participants will complete a shortened version of the Big 5 inventory. The Big 5 typically have weak associations with performance in lab studies of small-group problem solving (Bell, 2007). However, we believe that this very widely-used instrument may be helpful in understanding the characteristics of effective managers and team members. We will administer a variant of the 10-item version from Gosling.

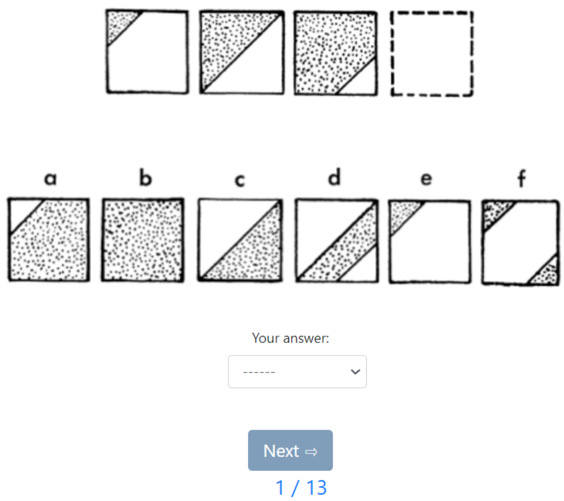
How strongly do you agree or disagree with the following statements?

Please indicate your answer on a 7 point scale from Strongly Disagree to Strongly Agree

	Disagree Strongly	Disagree Moderately	Disagree a little	Neither	Agree a little	Agree moderately	Strongly agree
I am talkative	○	○	○	○	○	○	○
I enjoy meeting new people	○	○	○	○	○	○	○
I am kind and considerate to others	○	○	○	○	○	○	○
I am helpful and cooperative	○	○	○	○	○	○	○
I am organized and efficient	○	○	○	○	○	○	○
I am hardworking and persistent	○	○	○	○	○	○	○
I am calm and emotionally stable	○	○	○	○	○	○	○
I am not easily upset	○	○	○	○	○	○	○
I am interested in new and different things	○	○	○	○	○	○	○
I am imaginative and creative	○	○	○	○	○	○	○

IQ (Culture Fair Intelligence Test)

Participants will complete two forms of the Culture Fair Intelligence Test (CFIT III). This is a well-known test of fluid intelligence. An example item is provided below. In the example, participants are asked to examine the pattern established by the top row of three boxes, and to indicate which of the options on the second row (a to e) best fits the pattern, and fills the dotted box on the top row.



2.3 Group measures

This section describes the group task we use to measure collaborative problem solving: the group allocation game. We sought a task that:

1. Has right and wrong answers, i.e. objective measures of performance
2. Requires collaborative problem solving, i.e. tasks in which working together is essential
3. Has analogous tasks that can be administered to individuals, to allow us to condition on individual task-specific skill
4. Has a clear role for the manager that includes elements of co-ordination, delegation and motivation.

2.3.1 Group Introduction

After groups have been allocated including the assignment of leadership, group members are informed of their role and given 60 seconds to introduce themselves. The manager is encouraged to facilitate this interaction.

2.3.2 Group Allocation Task

In this task, the goal for the group is to score as many points as possible by answering reasoning questions. The group is given 8 minutes. There are three question modules: numerical, analogical and spatial. For each correct answer the team scores 1 point and for each incorrect the score is reduced by 0.5 points. A group's score is the minimum module score. For example, if the module scores are as follows: 4, 8 and 9, then the total team score is 4. This scoring is similar to the weakest link coordination game so the effective group effort equals the lowest effort expended by an individual in the group. The weakest-link feature of this task resembles many real life situations where the performance of the entire group depends on the worst performer (e.g., Hirshleifer, 1983).

During the task, the manager is sitting in between the two 'workers' and is able to talk to them throughout the task. Every 2 minutes the task is paused for 60 seconds and there is an additional opportunity for the manager to communicate with their team. Managers are informed of their team's overall score, while workers are only informed of their individual module score.

The role of the manager is as follows:

- Assign workers to different modules. Managers can reassign team members at any point during the task.
- Monitor progress.
- Communicate with their team about the proposed strategy.
- Motivate the team.

We argue that this task mirrors key features of real-world management. In particular, our task requires:

1. Coordination

The Group Allocation task requires the manager to coordinate among group members to ensure the highest score possible. In particular, the score of this task depends not only on the actions of individuals but also on the actions of the whole team. Only the manager is aware of the team's score (i.e., the minimum score). But the manager is not aware of the module that requires attention (i.e. they don't know which module is the lowest score and needs more work). To find this out, managers must communicate with their team members, who know the score in the module they are working on. The task requires the manager to communicate and negotiate in order to coordinate their efforts, as the lack of coordination may lead to an inefficient outcome for the group as a whole. The manager must then coordinate with team members to ensure efficient group allocation and the highest score possible.

Coordination is vital to team outcomes. Coordination can be critical to resolving collective action problems and market failures. It has been shown to be important for conflict prevention, halting the transmission of viruses, changing social norms, improving firm performance, optimizing resource use on common land, and raising the provision of public goods. For instance, managers are able to coordinate their followers around a credible mission statement, which communicates the future course of action of the organization (Bolton et al, 2008).

2. Communication

This task requires the manager to regularly communicate with their group to make sure the group isn't pointless spending resources on a task where they already have a good score. The manager must also communicate with group members to discuss strategy and identify team member preferences etc. This is further facilitated by the pause and the fact the manager sits inbetween the other team members.

Good communication is often considered a crucial aspect of effective management skill. Managers who communicate effectively can articulate their vision and goals clearly, build trust and rapport with their team, and foster a culture of open communication and collaboration. In contrast, managers who struggle with communication may have difficulty conveying their ideas, fail to inspire their team, and struggle to maintain productive working relationships

3. Delegation/ strategizing

In this task one of the key roles of a manager is to delegate modules to team members. The manager has to process information such as the current team performance and well as individual module performances and must decide how to allocate group members across the modules.

A manager's ability to delegate is important because it allows them to effectively distribute tasks and responsibilities among team members based on their skills and strengths. This not only helps to ensure that tasks are completed efficiently and effectively to improve group performance. In a laboratory experiment, Fehr, Herz, and Wilkening (2013) have shown that people tend to be reluctant to give up decision rights, even if delegation is optimal from an organizational point of view, and often also in terms of own material well-being. Fehr et al. (2013) show that more than half of their participants are unwilling to delegate control to a subordinate even if doing so is profitable given what the participants themselves believe, increases own profits by 45% in the actual data, and increases overall profits for the organization by 30% in the actual data.

4. Motivation

Throughout this task the manager is able to communicate with group members, as part of this communication they are able to motivate team members to complete the somewhat repetitive task. For instance, during the task there are two pauses where the leader is able to motivate workers.

Motivating workers is important for a manager because it can lead to higher levels of job satisfaction, better performance, increased productivity, and overall success for the organization. A manager who can effectively motivate workers can also foster a positive work environment, build stronger relationships with team members, and create a culture of collaboration and innovation. Additionally, motivated workers are more likely to be engaged in their work and to remain committed to the organization, leading to lower turnover rates and reduced costs associated with hiring and training new employees.

2.4 Incentives and payments

The suggested incentive regime is summarized in the table below. The ranges for the Manager incentives reflect the difference between the highest and lowest performing Manager. We expect including phase 1 and phase 2 average earnings will be around £30-40. All payments will take place at the end of phase 2. Details are provided below.

Phase 1: Individual tests

- Subjects will receive a completion fee of that is equivalent to approximately £8/hour. We will base our time estimate on the pilot; we're aiming for around 30 minutes.

- Subjects will be paid for performance on the individual tests (Assignment Game; Emotional Perceptiveness; CFIT). The maximum bonus will be £4.

Phase 2: Group testing

- Subjects receive a show up fee that is equivalent to approximately £8/hour. We will base our time estimate on the pilot; we're aiming for between 90-120 minutes.
- 'Workers' will also receive a flat fee of £8 for participating in the group allocation task.

Each participant will be allocated to 4 different groups.

To provide an additional incentive to managers, we will randomly select 1 of the 4 group allocation rounds as a bonus round. Managers of groups in the top 40% of performers in the bonus round will be paid a bonus £12, for those groups in the bottom 40% the manager will receive £4. All other group managers will be paid £8. Team members pay stays fixed as described above. This will ensure that like in the wild, managers whose team perform well are paid a bonus amount. However, the average pay of Managers and workers will be kept roughly equivalent. If there's a large discrepancy, people who are allocated to the Team member role might disengage and feel as though the experiment is unfair.

3. Analysis Plan

3.1 Notation and models

Let individuals be indexed by $i = 1, \ldots, n_i$ and groups be indexed by $g = 1, \ldots, n_g$.

Let I_{ig} be an indicator of whether participant i is in group g. I_{ig} is a vector of length n_g, where:

$$I_{ig} = \begin{cases} 1 & if\ i\ is\ in\ g \\ 0 & otherwise \end{cases}$$

I_{ig} is determined through the randomization process described in Section 2.1.4.

Next, define two roles: Manager and Worker (labelled as team members in the actual experiment). Let M_{ig} be an indicator of whether participant i is the manager for group g. Similarly, let W_{ig} be an indicator of whether participant i is a worker in group g.

Analysis ignoring roles

First, consider a general causal effect that an individual may have on a team *regardless of role.*

- Teamplayer index = β_i; this is the average causal contribution individual i makes to team performance, conditional on $i's$ skill at the underlying task.

To estimate the 'teamplayer index', let X_i denote the performance of individual i on the individual cryptography task. Let G_g denote the performance of group g in the two group tasks.

Some groups may perform better purely because they have a greater endowment of task-specific skill. To remove this factor – and other potential confounds – consider a baseline model for the output of group g:

$$G_g = \gamma \sum_i X_i(M_{ig} + W_{ig}) + \boldsymbol{\delta Controls_g} + \epsilon_g \qquad (1a)$$
$$\epsilon_g \sim N(0, \sigma_G^2)$$

The term $\sum_i X_i(M_{ig} + W_{ig})$ measures group g's endowment of task-specific skill. $\boldsymbol{Controls_g}$ is a set of other potential confounds, including an indicator for whether group g contain participants who were acquaintances, friends or colleagues outside the context of the experiment.

Notice that $\hat{\epsilon}_g$ can be thought of as a measure of *group level* efficiency. With only a single randomization, it is impossible to determine whether variation in ϵ_g arises from unmeasured individual attributes of particular team members, or from group dynamics between team members. Observing individuals in only one group would not allow for β_i to be estimated. However, by randomly assigning individuals to multiple groups we can assess whether $\hat{\epsilon}_g$ is correlated for individuals:

$$\hat{\beta}_i = \frac{1}{\sum_g(M_{ig} + W_{ig})} \sum_g \hat{\epsilon}_g \, (M_{ig} + W_{ig}) \quad (2a)$$

In our framework (see Weidmann and Deming, 2021) $\hat{\beta}_i$ is an estimate of the causal contribution of individual i to team performance. With only five team assignments $\hat{\beta}_i$ is relatively noisy at the individual level. Thus, we focus on σ_β, the standard deviation of the β estimates. We estimate σ_β using a multilevel model:

$$\hat{\epsilon}_{gi} = \beta_i + e_{gi} \quad (3a)$$
$$\beta_i \sim N(0, \sigma_\beta^2)$$
$$e_{gi} \sim N(0, \sigma^2)$$

Incorporating roles into the analysis

Next, consider **role-specific** causal effects an individual might have on a group:

- Manager index = α_i; this is the contribution individual i makes when they are in a manager role, independent of $i's$ skill at the underlying task.

- Worker index = Ω_i; this is the contribution individual i makes when they are in a team member role, independent of $i's$ skill at the underlying task.

We hypothesize that the task-specific skill level of Managers and Workers might affect group performance differently. Our preferred model is an extension of model 1a that allows for this possibility i.e.:

$$G_g = \gamma_\alpha \sum_i X_i M_{ig} + \gamma_\Omega \sum_i X_i W_{ig} + \boldsymbol{\delta Controls_g} + \epsilon_g \qquad (1b)$$

$$\epsilon_g \sim N(0, \sigma_G^2)$$

For each participant who assumes the role of a manager, we estimate the manager index (α_i) as the average $\hat{\epsilon}_g$ across their groups. We use the same approach for the worker index (Ω_i):

$$\hat{\alpha}_i = \frac{1}{\sum_g M_{ig}} \sum_g M_{ig}\, \hat{\epsilon}_g \quad (2b')$$

$$\hat{\Omega}_i = \frac{1}{\sum_g W_{ig}} \sum_g W_{ig}\, \hat{\epsilon}_g \quad (2b'')$$

Now, $\hat{\alpha}_i$ is an estimate of the causal contribution of individual $i's$ management skills. With enough randomizations, we could precisely estimate α_i for each participant who was allocated to the role of manager. Equivalent logic applies to workers and Ω_i. Once again, we are also interested in the spread of the α and Ω estimates. We will estimate σ_α and σ_Ω using a multilevel model:

$$\hat{\epsilon}_{ig} = \alpha_i M_{ig} + \Omega_i W_{ig} + e_{ig} \quad (3b)$$

$$\begin{pmatrix} \alpha_i \\ \Omega_i \end{pmatrix} \sim N\left(\mathbf{0}, \begin{matrix} \sigma_\alpha^2 & 0 \\ 0 & \sigma_\Omega^2 \end{matrix}\right)$$

$$e_{ig} \sim N(0, \sigma^2)$$

In model (4) $\hat{\epsilon}_{ig}$ is a vector of skill-adjusted group performance (1x$3n_g$), α_i is a random manager effect for individual i on group g, Ω_i is a random worker effect for individual i on group g, and e_{ig} is residual error.

$\hat{\sigma}_\alpha$ is the typical "manager effect": i.e. the impact on group performance of having a manager who is 1sd above average in terms of their management skills. Similarly, $\hat{\sigma}_\Omega$ is the typical "worker effect".

3.2 Inference

For inference on the key variance parameters $(\sigma_\beta, \sigma_a, \sigma_\Omega)$ we will follow Weidmann and Deming (2021) and report p-values from three different methods: randomization inference; inverted profile likelihood confidence intervals; and a simple wald test.

The randomization inference will proceed in 4 steps (using σ_β as an example):

First, we will control for individual skill, by fitting model (1a). Second, we simulate five thousand allocations of individuals to groups. Third, we fit model (3a) for each simulation and estimate $\sigma^2_{\beta(NULL)}$. Fourth, we compare our observed team player effect to the simulated distribution under the null, calculating how often the null distribution provides a more extreme value than $\hat{\sigma}_\beta$, i.e. $P(\sigma_{\beta(NULL)} > \hat{\sigma}_\beta)$. This is our p-value (Ernst, 2004).

3.3 Validation

We will test the predictive power of our manager/worker estimates using Leave One Out validation. By the end of the experiment each participant will have completed 4 group tasks. Our approach will proceed as follows:

1. Remove all data from group t
2. Estimate $\hat{\alpha}_i^{LOO}$ and $\widehat{\Omega_\iota}^{LOO}$ using models (2b') and (2b'')
3. Fit the following model, using the held-out data from task t

$$G_g = \gamma_{\alpha^{\text{LOO}}}^t \sum_i \hat{\alpha}_i^{LOO} L_{ig} + \gamma_\Omega^t \sum_i \widehat{\Omega_\iota}^{LOO} F_{ig} + \epsilon_g \qquad (LOO_t)$$

4. Return to step 1, until we have worked through $t \in \{1, \dots, 5\}$

We will then report the average coefficients $\gamma_{a^{Loo}} = \Sigma_{t=1}^5 \gamma_{\alpha^{\text{Loo}}}^t$ and $\gamma_{\Omega^{Loo}} = \Sigma_{t=1}^5 \gamma_{\Omega^{\text{Loo}}}^t$

3.4 Analysis by research question

RQ1: Can we identify good managers? If so, what skills do they have?

a. Can some managers reliably improve team performance (conditional on task-specific skills of team members)?

   Do we reject the null that $\hat{\sigma}_a = 0$? We will use the point estimate from model 3b and conduct hypothesis tests using randomization inference [with profile likelihood and Wald tests for robustness]. To answer RQ1a, we will also report on the LOO validation described in section 3.3.

b. What characteristics, skills and traits are associated with good managers?

   We will calculate marginal associations between $\hat{\alpha}_i$ and: emotional perceptiveness (RMET); political savvy (Political Skill Inventory); Indecisiveness (II); personality (Big5); and IQ (CFIT). In each case we will present simple scatterplots and examine non-linear relationships using cubic splines. We will also examine whether demographic factors (age, gender) are associated with $\hat{\alpha}_i$.

c. How important are management skills compared to task-specific skills?

   We will report point estimates of $\hat{\sigma}_\alpha$ from model 3b. We will compare this to the effect of task-specific skill for managers ($\hat{\gamma}_a$ from equation 1b).

d. Do good managers have different skills from good workers?

   To address this question, we fit the following linear model:

$$\hat{\beta}_i = \gamma_{1i} AS_i + \delta_{1i} AS_i M_i + \gamma_{2i} SS_i + \delta_{2i} SS_i M_i + e_i$$

Where AS_i is a measure of allocative skill (combining the Assignment Game and the Indecision Index), and SS_i is a measure of social skill (combining the emotional perceptiveness test, and the Political Skill Inventory). Recall that β_i is the average causal contribution that people make to a group independent of their role. After fitting this model, we will test two hypotheses separately:

$H_0: \delta_{1i} = 0$ *and*

$H_0: \delta_{2i} = 0.$

RQ2: How should managers be selected?

a. Do people who nominate themselves for management roles have better management skills?

   We will report on the difference between the mean estimated manager index ($\hat{\alpha}_i$) for people who self-selected into the role and the mean manager index for people who were randomized into the manager role. We will also conduct a two-tailed null-hypothesis test. $E[\alpha_i|Z_1 = 1, M_i = 1] = E[\alpha_i|Z_1 = 0, M_i = 1].$

b. What is the relationship between gender, management skills, and self-nomination?

   We will fit a model for the estimated manager index: $\hat{\alpha}_i \sim \gamma_0 + \gamma_1 Female_i + \gamma_2 Z_{1i} + \gamma_3 Female_i Z_{1i}$, and report all coefficients. This will let us examine whether there is a difference between men and women in 'self-knowledge' (i.e. being aware that one is a good manager and being willing to nominate oneself for the role). We can measure this by comparing the average manager index for self-selected vs random managers. For women, this quantity is given by $\gamma_2 + \gamma_3$. For men the equivalent quantity is γ_2. We answer the question of a possible gender difference by testing the null hypothesis of $\gamma_3 = 0$.

c. How do the managerial skills of the top-third (and bottom-third) of people who nominate themselves compare with the managerial skills of the top-third of managers that emerge from randomized assignment
   To study this we compare the two distributions.

d. Do managers have a good understanding of the skills required to be a good manager? Do workers?
   We ask participants to rank the skills they believe are required for managers to be successful in our experiment (allocative skill; social skill; IQ; task skills). For each skill, we will separately calculate the average rank for workers and for managers. We will then compare these ranks to the ground truth, which will be generated by listing each skill in order of its association with $\hat{\alpha}_i$. To mitigate issues with measurement error, we will use disattenuated point estimates of correlation coefficients.

3.5 Additional exploratory analyses

We will investigate the effect that managers have on specific individuals…

- we have individual measures of proficiency at numbers/analogies/spatial (henceforth 'task skills')
- conditional on task skills, do individuals (rather than groups) under or overperform when they're working for a randomly assigned manager?
- the intuition is as follows: when I'm working for a good/bad manager, do I do better/worse on tasks than I did by myself, even though the manager isn't directly helping? If yes, then there's evidence that the manager is improving/hurting my performance (probably through motivation)
- we will look at questions like:
    - 'do female managers help improve female worker performance';
    - ‘do some managers excel at motivating individuals who scored poorly on the individual tasks’
    - We also study learning and whether good managers show a steeper learning curve.

References

Adhvarya, A., Gunnsteinsson, S., Murathanoglu, E. & Nyshadham, A. (2021). (Mis)allocation of managerial training within the firm. Mimeo: Duke University.

Ahearn, K. K., Ferris, G. R., Hochwarter, W. A., Douglas, C., & Ammeter, A. P. (2004). Leader political skill and team performance. *Journal of management, 30*(3), 309-327.

Anderson, C.R. & Schneier, C. E (1978). Locus of control, leader behavior & leader performance among management students. The Academy of Management Journal, 21(4), 690-698.

Arora, A. (2022). Election by Community Consensus: Effects on Political Selection & Governance. The Review of Economics & Statistics, 104 (2), 321–335.

Austin, J. R. (2003). Transactive memory in organizational groups: the effects of content, consensus, specialization, and accuracy on group performance. *Journal of applied psychology, 88*(5), 866.

Baddeley, A. D. (2001). Is working memory still working? *American psychologist, 56*(11), 851.

Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., & Plumb, I. (2001). The "Reading the Mind in the Eyes" test revised version: A study with normal adults, and adults with Asperger syndrome or high-functioning autism. *Journal of child psychology and psychiatry, 42*(2), 241-251.

Bass, B. M., & Avolio, B. J. (1992). Multifactor leadership questionnaire-short form 6S. *Binghamton, NY: Center for Leadership Studies*.

Beenen, G., Pichler, S., Livingston, B. and Riggio, R., 2021. The good manager: Development and validation of the managerial interpersonal skills scale. *Frontiers in psychology*, *12*, p.631390.

Bell, S. T. (2007). Deep-level composition variables as predictors of team performance: a meta-analysis. *Journal of Applied Psychology, 92*(3), 595.

Bertrand, M. and Schoar, A., 2003. Managing with style: The effect of managers on firm policies. *The Quarterly journal of economics*, *118*(4), pp.1169-1208.

Blau, F.D., Ferber, M., Winkler, A. (2002). The Economics of Women, Men & Work. 4th edition. Englewood Cliffs, NJ: Prentice Hall

Brandt, J. (1991). The Hopkins Verbal Learning Test: Development of a new memory test with six equivalent forms. *The Clinical Neuropsychologist, 5*(2), 125-142.

Brunell, A. B., Gentry, W. A., Campbell, W. K., Hoffman, B. J., Kuhnert, K. W., & DeMarree, K. G. (2008). Leader emergence: The case of the narcissistic leader. Personality & Social Psychology Bulletin, 34, 1-14.

Chabris, C. F. (2007). Cognitive and neurobiological mechanisms of the Law of General Intelligence. In M. J. Roberts (Ed.), *Integrating the mind: Domain general vs domain specific processes in higher cognition* (pp. 449–491). Psychology Press.

Chakraborty, P. & Serra, D. (2021). Gender & leadership in organizations: Promotions, demotions, & angry workers. Working Papers 20210104-001, Texas A&M University, Department of Economics.

Colom, R., Rebollo, I., Abad, F. J., & Shih, P. C. (2006). Complex span tasks, simple span tasks, and cognitive abilities: A reanalysis of key studies. *Memory & Cognition, 34*(1), 158-171.

Davis, M. H. (1983). Measuring individual differences in empathy: Evidence for a multidimensional approach. *Journal of personality and social psychology, 44*(1), 113.

Dennett, H. W., McKone, E., Tavashmi, R., Hall, A., Pidcock, M., Edwards, M., & Duchaine, B. (2012). The Cambridge Car Memory Test: A task matched in format to the Cambridge

Face Memory Test, with norms, reliability, sex differences, dissociations from face memory, and expertise effects. *Behavior Research Methods, 44*(2), 587-605.
Desseranno, E., Stryjan, M., & Sulaiman, M. (2019). Leader selection & service delivery in community groups. American Economic Journal: Applied Economics, 11(4), 240-267.
Duchaine, B., & Nakayama, K. (2006). The Cambridge Face Memory Test: Results for neurologically intact individuals and an investigation of its validity using inverted face stimuli and prosopagnosic participants. *Neuropsychologia, 44*(4), 576-585.
Eagly, A.H., Karau, S.J., 2002. Role congruity theory of prejudice toward female managers. Psychological Review. 109(3), 573-597.
Edmondson, A. (1999). Psychological safety and learning behavior in work teams. *Administrative Science Quarterly, 44*(2), 350-383.
Engel, D., Woolley, A. W., Jing, L. X., Chabris, C. F., & Malone, T. W. (2014). Reading the mind in the eyes or reading between the lines? Theory of mind predicts collective intelligence equally well online and face-to-face. *PloS one, 9*(12), e115212.
Ertac, S. & Gurdal, M.Y. (2012). Personality, group decision-making & leadership. Koç University-TUSIAD Economic Research Forum Working Papers 1227.
Ferris, G. R., Treadway, D. C., Kolodinsky, R. W., Hochwarter, W. A., Kacmar, C. J., Douglas, C., & Frink, D. D. (2005). Development and validation of the political skill inventory. *Journal of management, 31*(1), 126-152.
Fiske, S. T., Cuddy, A. J., & Glick, P. (2007). Universal dimensions of social cognition: Warmth and competence. *Trends in cognitive sciences, 11*(2), 77-83.
Frost, R. O., & Shows, D. L. (1993). The nature and measurement of compulsive indecisiveness. Behaviour Research and Therapy, 31(7), 683-692.
Goldberg, L. R. (1992). The development of markers for the Big-Five factor structure. *Psychological assessment, 4*(1), 26.
Goldman, M. & Fraas, L. A. (1965). The effects of leader selection on group performance. Sociometry, 28(1), 82-88.
Gruenfeld, D.H., Inesi, M.E., Magee, J.C. & Galinsky, A.D. (2008). Power & the objectification of social targets. Journal of Personality & Social Psychology. 95, 111-127.
Haier, R. J., Colom, R., Schroeder, D. H., Condon, C. A., Tang, C., Eaves, E., & Head, K. (2009). Gray matter and intelligence factors: Is there a neuro- *g* ? *Intelligence, 37*(2), 136–144
Hollander, E.P., Fallon, B.J., ad Edwards, M.T. (1977). Some aspects of influence & acceptability for appointed & elected group managers. The Journal of Psychology, 95, 289-296.
Hollingshead, A. B. (1998). Groupand Individual Training: The Impact of Practice on Performance. *Small Group Research, 29*(2), 254-280.
Ingoglia, S., Lo Coco, A., & Albiero, P. (2016). Development of a brief form of the Interpersonal Reactivity Index (B–IRI). *Journal of Personality Assessment, 98*(5), 461-471.
IPIP. (2018). Retrieved from https://ipip.ori.org/New_IPIP-50-item-scale.htm
Kahneman, D. and Klein, G., 2009. Conditions for intuitive expertise: a failure to disagree. *American psychologist*, *64*(6), p.515.
Kane, M. J., Hambrick, D. Z., Tuholski, S. W., Wilhelm, O., Payne, T. W., & Engle, R. W. (2004). The generality of working memory capacity: a latent-variable approach to verbal and visuospatial memory span and reasoning. *Journal of Experimental Psychology: General, 133*(2), 189.

Kim, H.T., 2020. Linking managers' emotional intelligence, cognitive ability and firm performance: Insights from Vietnamese firms. *Cogent Business & Management*, *7*(1), p.1829272.

Knight, P.A. & Weiss, H. M. (1980). Effects of selection agent & leader origin on leader influence & group member perceptions. Organizational Behavior & Human Performance, 26, 7-21.

Kosfeld, M. & Rustagi, D. (2015). Leader punishment & coopertion in groups: Experimental field experiment from commons management in Ethiopia. American Economic Review, 105(2), 747-783.

Larson, J. R. (2010). *In search of synergy in small group performance*: Psychology Press.

Laughlin, P. R., Bonner, B. L., & Miner, A. G. (2002). Groups perform better than the best individuals on letters-to-numbers problems. *Organizational Behavior and Human Decision Processes, 88*(2), 605-620.

Laughlin, P. R., Hatch, E. C., Silver, J. S., & Boh, L. (2006). Groups perform better than the best individuals on letters-to-numbers problems: effects of group size. *Journal of Personality and social Psychology, 90*(4), 644.

Martin, S.L. & Campbell, H.L.E.M. (2013). Directive versus empowering leadership: A field experiment comparing impacts on task proficiency & proactivity. The Academy of Management Journal, 56(5), 1372-1395.

Minni, V., 2022. Making the Invisible Hand Visible: Managers and the Allocation of Workers to Jobs. *Job Market Paper*.

Mitchell, T.R., Smyser, C.M. & Weed. S.E. (1975). Locus of control: Supervision & work satisfaction. The Academy of Management Journal, 18, 623-630.

Nevicka, B., De Hoogh, A.H.B., Van Vianen, A.E.M., Beersma, B. & McIlwain, D. (2011). All I need is a stage to sine: Narcissist's leader emergence & performance. The Leadership Quarterly, 22, 910-925.

Nevicka, B., Ten Velden, F.S., De Hoogh, A.H.B. & Van Vianen, A.E.M. (2011). Reality at odds with perceptions: Narcissistic managers & group performance. Psychological Science, 22(10), 1259-1264.

Nisbett, R. E., Aronson, J., Blair, C., Dickens, W., Flynn, J., Halpern, D. F., & Turkheimer, E. (2012). Intelligence: new findings and theoretical developments. *American psychologist, 67*(2), 130.

Pinkham, A. E., Penn, D. L., Green, M. F., Buck, B., Healey, K., & Harvey, P. D. (2013). The social cognition psychometric evaluation study: results of the expert survey and RAND panel. *Schizophrenia bulletin, 40*(4), 813-823.

Reuben, E.M, Rey-Biel, P., Sapienza, P. & Zingales, L. (2012). The emergence of male leadership in competitive environments. Journal of Economic Behavior & Organization, 83(1), 111-117.

See, K.E., Morrison, E.W., Rothman, N.B. & Soll, J.B. (2011). The detrimental effect of power on confidence, advice taking & accuracy. Organizational Behavior & Human Decision Processes. 116: 272-285.

Sullivan, K. (2005). Alternate forms of prose passages for the assessment of auditory–verbal memory. *Archives of clinical neuropsychology, 20*(6), 745-753.

Tost, L.P., Gino, F. & Larrick, R.P. (2013). When power makes others speechless: The negative impact of leader power on team performance. The Academy of Management Journal, 56(5), 1465-1486.

Van der Vegt, G.S., de Jon, S.B., Bunderson, J.S. & Molleman, E. (2010). Power asymmetry & learning in teams: The moderating role of performance feedback. Organization Science, 21, 347-361.
Wegner, D. M. (1987). Transactive memory: A contemporary analysis of the group mind. In *Theories of group behavior* (pp. 185-208): Springer.
Weidmann, B., & Deming, D. J. (2020). *Team Players: How Social Skills Improve Group Performance* (0898-2937). National Bureau of Economic Research.
Woolley, A. W., Chabris, C. F., Pentland, A., Hashmi, N., & Malone, T. W. (2010). Evidence for a collective intelligence factor in the performance of human groups. *Science, 330*(6004), 686-688.