# Task Documentation

## General note

For all the tasks, we’d like the groups to submit a single answer. But, in some of the tasks we will ask individuals to work separately. For example, in the optimisation task, each person gets [6] guesses; i.e. there are 3 laptops in the room (1 per person) and each one allows a maximum of 6 guesses. At the end of the task, the group comes up with a single answer for where the global maximum is.

## Overview of Tasks

1. *Optimisation Task.* This is a task that hasn’t been used before. But we’re hoping for it to look as follows. Participants get the following instructions:

Your task is to try to find the number that generates the biggest possible output. You each have **[6]** guesses, which you can enter into your own laptop. A guess can be any number between [0] and [300]. After you enter your guess, the computer will give you back a number. There is a systematic relationship between the number you guess, and the number you receive, but the relationship will often be difficult for you to understand. Every time you type in the same number, the number you receive will be similar (but there may also be a bit of randomness added in). Usually, numbers that are close to each other will receive outputs.

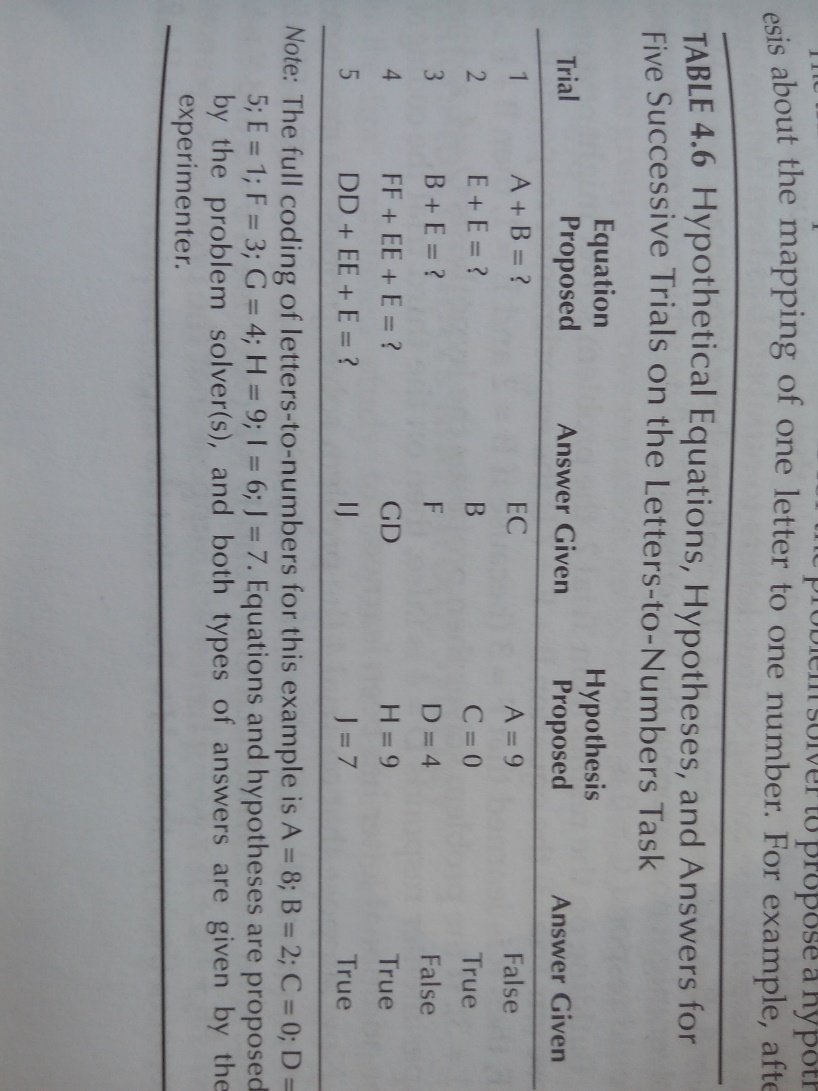
We will have underlying functions, which will be some sort of polynomial, with noise. Ideally we’ll have a time limit for each guess. i.e. every 20 seconds individuals are allowed to make another guess [so the total guessing period is 120 seconds]. At the end, groups have to decide what their collective “best guess” is.

1. *Cryptography Task.*

Before the task begins, the letters A-J are randomly mapped to the numbers 0-9. The goal for participants is to decipher this mapping in the minimum number of “trials”.

A trial involves three steps:

1. *Propose an Equation:* i.e. the group nominates the left-hand side of an equation, using letters, addition and subtraction: e.g. “A+B”
   * The group then receives an answer, e.g. “A+B=EC”
2. *Propose a Hypothesis*: the group makes a guess as to one element of the mapping, e.g. “E=1”
   * The group then gets confirmation about whether their guess is correct: e.g. “E=1 is TRUE”
3. *Guess full mapping:* at the end of each trial, the group guesses at the whole mapping. If they are correct, the task is complete. Otherwise, they start a new trial. We will probably limit the number of trials [e.g. to 10 trials].



1. *Brainstorming.* The core idea is that we provide a prompt, and then participants come up with as many ideas as possible. Some examples:
   1. *“List as many words as possible that start with S and end in N.*”
   2. *“List different uses of a brick.*”
   3. *“List as many equations as as possible that equal 10 while using the operators +,–,/,\** *and using the numbers between 2 and 8 only once per equation.*”

Some of these can be scored automatically (e.g. a and c). Some will be scored manually (e.g. b). In both cases we need to identify unique answers. Specifically, we’d like to award 1 point for each distinct “correct answer” a group comes up with, plus up to an additional point if the answer is something no other group came up with. This is a task where we’d like everyone to be able to enter their answers separately, and then have the computer aggregate the group’s response into a long list.

1. *Unscramble Words.* We give groups a list of 24 words with all the letters scrambled. They have two minutes to unscramble as many words as possible. Each item on the list had only one correct answer, and groups received one point for each correct answer.
2. *Group Matrix Reasoning.* Raven’s Advanced Progressive Matrices (RAPM) is a standardized test of general fluid reasoning capacity. [We need to see about getting licenses for this!]. Each item presents a 3 x 3 array (or “matrix”) of items with the lower-right corner empty. The test is then multiple choice: based on patterns in the array, subjects must pick which of eight answer pieces properly belongs in the empty space. There is only one correct answer per question.
3. *Judgment.*
   1. We present groups with 5 pieces of art. Ideally, we’d have them presented sequentially, and then all shown together at the end. We ask groups to judge, on a scale of 1 (ugly) to 10 (beautiful) how a large sample of Americans rated them. [We establish the ground truth with a survey that we’ll conduct separately].
   2. For the second judgment task, groups were asked to judge the number of pages in books based on images of their page edges. We present an image of a book (taken from side on). Groups enter a number. Groups receive points proportional to the number of book pages that their estimate erred, corrected for the total number of pages in the book (i.e. 10 pages wrong out of a 400-page book is a smaller error than 10 pages wrong out of a 50-page book).
4. *Memory.* The battery includes two memory tasks.
   1. The first one is a video memory task in which group members are shown a 90-second video twice and then asked a set of (multiple-choice) questions about its features afterwards. They received points for each correct answer.
   2. The second task was exactly the same as the first one, except that group members are shown a complex image instead and answered (multiple-choice) questioned about its features.
5. *Detection.* The idea here is that groups are presented with either a large grid of words or with a large grid of images. In each case, they are asked to judge the characteristics of the grid: “What is the most frequent object in the grid?”, “Which is the least frequent object in the grid?”, and “Which two objects have the most similar frequency?”. The time limit is 60 seconds per grid in order to make it impossible for any individual member to exhaustively count all objects in the grid. Groups receive points for the correct answer and partial credit based on how close their answers were to the correct answers.
6. *Decision-making task [THIS IS A WORK IN PROGRESS]*

The core idea is that each participant is given a slightly different set of information about a problem. For example, a company is trying to decide between investments A, B, or C. Each has strengths and weaknesses. Each of the three participants receives “common information”, which are facts that everyone gets, and also “individual information” which only one member of the group receives. We’ll give them [5] minutes to read through the information in silence, then we’ll take the info away and ask them to discuss the problem, and come to a collective decision. The challenge, for the group, is to surface the “individual information” so that they make a decision based on all available info (rather than just the common information). So, the technical challenge here is that, in a group of 3 people, we’d need each person to see different information, for a fixed period of time.­­­