



# **Questionnaire, Validation and Testing**

# C-me background and development

## Background

The purpose of the C-me report is to help individuals better understand themselves, recognise the difference in others' behavioural preferences and to use this knowledge to improve relationships and interactions.

To succeed in this aim, the report has to be based on sound science; both in its psychological basis and in the correct measurement of the elements used in order to produce the reports.

The aspects of personality that drive behaviour need to be consistently measurable when individuals score themselves, as well as observable by others. This measure then needs to be presented in such a way that individuals can clearly recognise themselves in the results and be able to use the results for personal development.

The key elements are:

- The measurement of preferences through the questionnaire
- The expression of those preferences in a way that resonates with individuals and is helpful to others.

Critical in our testing is therefore the consistency of the questionnaire output and the individual resonance of the report statements.

C-me was shaped after looking at and working with dozens of 'personality' systems and observing behaviour in the workplace and on training programmes. The critical filter was our desire for a system that would lead to real impact - opportunities for personal development and clarity for individuals to understand themselves and others. C-me has now produced tens of thousands of reports, and this remains our ongoing filter for development and honing.

The language we use in the reports arises from the combination of four of Jung's preferences; his orientations, Extraversion and Introversion and his attitudes, Thinking and Feeling. These preferences are used to underpin the majority of modern psychometrics. We want to make them accessible, memorable and non-binary. We found the best experience for clients came when they were mapped on to the language of colour. This brings the preferences to life and gives a simple and memorable way to communicate them and apply them in everyday life.

## The building blocks of our measures (in Jungian terms) are:

**RED:** Extraverted Thinking.

Demanding, competitive, strong willed, bold, driving, challenging, determined, forceful.

**YELLOW:** Extraverted Feeling.

Engaging, friendly, enthusiastic, optimistic, spontaneous, talkative, open, sociable, informal.

**GREEN:** Introverted Feeling.

Empathetic, patient, cooperative, helpful, encouraging, considerate, supportive, diplomatic.

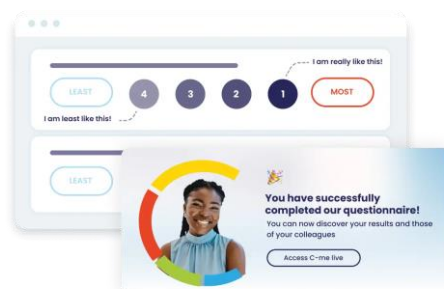
**BLUE:** Introverted Thinking.

Detailed, particular, questioning, logical, reflective, precise, thorough, reasoned, consistent.

We set out to measure how much of each of these four energies an individual recognises in themselves.

## Questionnaire

In our questionnaire we really only ask one question: “How much of each colour energy do you see in yourself?” We ask multiple questions to be sure of the end result.

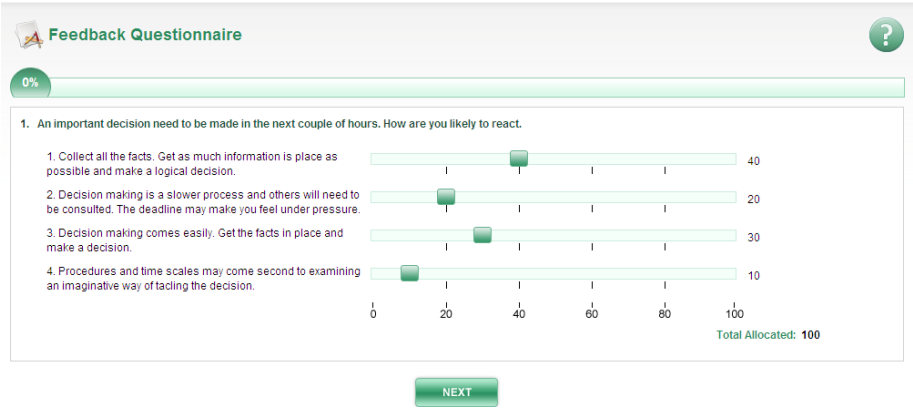


We chose sets of three words to describe each aspect of behaviour. (Any of these words in isolation could be linked to more than one colour, but by grouping in threes we build up a more accurate picture of different energies that an individual may identify with.) Each question has four sets of three words.

We ask individuals to rank these four sets of words 1 (most like them) 2, 3, and 4 (least like them). A weighting is assigned to each choice, and the algorithm attributes this to a colour, Red, Yellow, Green, Blue. These weightings are averaged over the 15 questions and expressed on the Adapted Behaviour graph as a percentage.

Statistically, 12 questions give a valid result, but we retain 15 questions to allow us to research new or better questions.

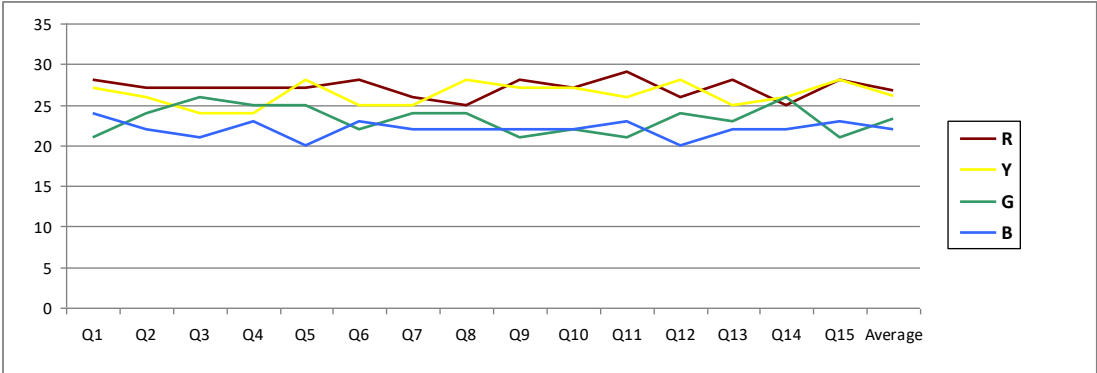
During the development phase we experimented with several alternative methods of asking questions. For example, we tried a slider mechanism and situational questions. However, our word sets gave the best validation results.



**Validation**

Once we had a questionnaire that feedback suggested consistently measured the relative strengths of the Red, Yellow, Green and Blue energy we then began testing more widely amongst varied client cohorts. C-me is unusual in being tested from the outset with representative individuals testing in the context of real team application rather than test groups who would have self-selected and tested in abstract situations.

Here is a typical data export of 200 results showing the responses to each question.



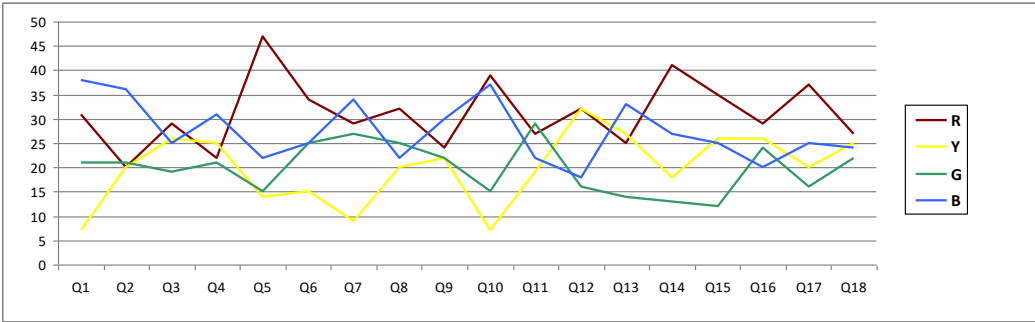
No individual question produces a result that is significantly different from the overall average.

To “improve” the questionnaire in this output we may need to make the Green choice in question 3 slightly less desirable and the Red choice slightly more desirable. In this way the individual question average would be closer to the overall average. In early testing we wanted to achieve a change in any anomalies in one in ten respondents.

The exact word choices are critical and often small differences give big results.

Questions with results that fall outside a narrow band were rewritten or deleted. Once a question is altered the testing needed to be repeated.

Compare the results above with an early test of the slider questionnaire below.

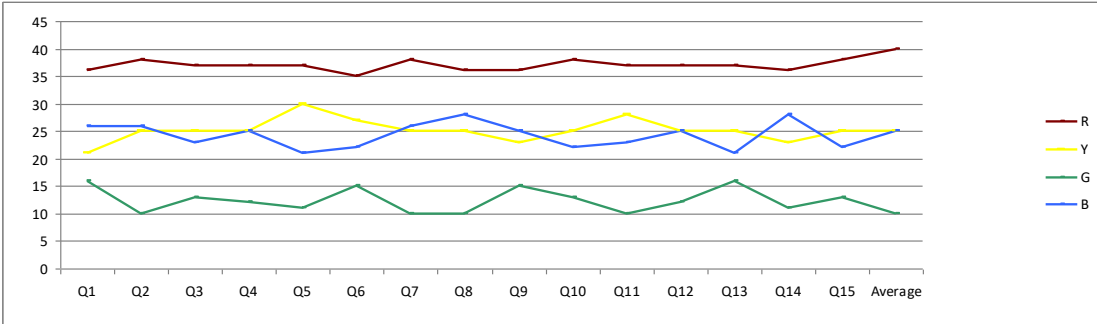


The sample size is much smaller (42) but the results are inconsistent and showed that we were heading in the wrong direction.

**Two other elements of the validation process were:**

**1) Individual Colours testing**

**Here are 10 participants with an average for Red preference of over 35%:**



We used this type of selection because this was a sample that had a clear preference and they are telling us about the strength of this preference.

We can see that each individual question also has a clear Red preference.

Where Red is the highest score, Green is the lowest score. Psychologically Red and Green are opposites (Extraverted Thinking/Introverted Feeling). The results showed that the questionnaire reflects this.

This exercise was repeated for high Yellow scores, high Green and high Blue with the same consistent results.

## 2) Split half testing

	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Average
<b>R</b>	26	24	23	25	27	24	24	22	24.375
<b>Y</b>	27	26	25	26	26	29	27	26	26.5
<b>G</b>	24	26	28	25	22	25	23	25	24.75
<b>B</b>	22	22	21	22	23	20	24	24	22.25
	Q9	Q10	Q11	Q12	Q13	Q14	Q15	Q16	Average
<b>R</b>	22	28	23	27	23	24	23	25	24.375
<b>Y</b>	25	24	28	27	28	26	27	27	26.5
<b>G</b>	29	20	24	22	27	25	27	23	24.625
<b>B</b>	22	26	23	22	20	23	21	23	22.5

From this sample you can see that questions 1-8 results are indistinguishable to questions 9-16; which is a good test of the integrity of the questionnaire (the original questionnaire had 16 questions but on further testing and validation this was further refined leaving the current version with 15 questions).

## Graphs

Each profile report displays two graphs; the Adapted Behaviour and the Natural Behaviour.

### The Adapted Behaviour

From the overall average of the scores of the fifteen questions we produce the Adapted Behaviour graph; how the individual sees themselves.

The relative strengths of the colour energies on this graph are used to select statements from the thousands of statements in the database that produce the reports.



The Adapted Behaviour graph in the report is the average RYGB scores from the questionnaire.

In this example, for almost every question the participant ranked the corresponding Red or Blue choices highest (1 or 2). Green was occasionally chosen for a higher score but Yellow was almost always chosen as the lowest score, average 11%.

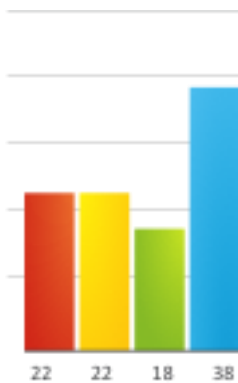
Selecting one colour as 'most like' for every question would give a score of 40%

Selecting one colour as 'least like' every question would give a score of 10%.

There are no good or bad graphs, no good or bad scores and no good or bad profiles. What we are doing in this graph is reporting that this is the result of what the individual has told us about their preferences. Is it 'true'? Perhaps. Is this the whole person? No. It probably gives us some excellent clues as to how they see themselves and we should ask them to hone the information and expand upon it.

## The Natural Behaviour

The Natural Behaviour graph is a representation of an individual's less conscious or 'natural' behaviour. This can often be their blind spot as it reports their unmodified behaviours. In reporting on the colour energies in the Adapted Behaviour graph we describe their modified behaviour as "who we would like to be, or how we adapt and modify our behaviours to be a success in the context we're in". This graph seeks to give an insight into how an individual acts when they are less conscious of their behaviour. This can be when they are under pressure or also when they are feeling very relaxed.



This graph disregards the top two choices as *conscious* choices of who we would like to be, focusing on the unconscious allocation of ranks 3 and 4 (least like the individual). We then use the information that Red and Green are psychological opposites. Blue and Yellow are psychological opposites. Saying 'I am least Yellow' is therefore generally unconsciously expressing a preference for Blue.

This Natural Behaviour graph, indirectly, gives another view of the personality that may give us a better insight into less conscious behaviour.

Each individual is a unique blend of the 4 colour energies that makes them the person that they are. In this representation of the natural behaviour their preference is concentrated in a smaller range of behaviour that is more likely to come to the fore when they 'revert to type'.

**An example of a Natural Behaviour score calculation**

Word sets are ranked 1-4, 1 being most like and 4 being least like. Each ranking carries a corresponding weighting:

- Ranked 1      Weighting 4
- Ranked 2      Weighting 3
- Ranked 3      Weighting 2
- Ranked 4      Weighting 1

If an individual question is ranked R 1, Y 2, G 3, B 4.

The 1 & 2 rankings are ignored.

By ranking B as 4, the system gives Y a new ranking of 1, and a weighting of 2 (saying I am least like Blue can be taken as saying I am most like Yellow).

By ranking G as 3, the system gives R a new ranking of 2, and a weighting of 1 (saying I am least like Green can be taken as saying I am most like Red).

	<b>Ranking</b>	<b>Weighting (Adapted Behaviour)</b>	<b>Weighting (Natural Behaviour)</b>
<b>R</b>	1	4	1
<b>Y</b>	2	3	2
<b>G</b>	3	2	0
<b>B</b>	4	1	0

The resulting score can have a maximum of 66%. We use a graph with a scale of 55% to match the left graph. This decision followed research showing that differing scales led to confusion and misinterpretation of the results, for example the degree of modification. This is a compromise between a better graphical representation and a correct mathematical presentation. Our apologies to those with strong Blue preference!



The individual's greatest potential can often lie in recognising and controlling the behaviours that can emerge from this less conscious or 'natural' behaviour.

## Cross types

Approximately 5-10% of reports completed result in graphs that show opposite energies as the two highest choices. We call this a 'cross type' or 'creative type'.



Whilst more uncommon this can be a creative and dynamic combination. It can, however, also mean that colleagues may find it a little more difficult to read an individual's behaviour as it changes between contrasting energies.

With this combination there is a higher chance that the results may have been affected by external circumstances around completion.

In this example the difference between the top 3 colours is marginal and consequently the accuracy of the resulting report is delivered with less certainty.

A more pronounced graph pattern may give a more accurate picture.

Where a report doesn't strongly resonate we would recommend the individual repeat the questionnaire.

Approximately half of the individuals displaying cross type preferences will change to a more usual preference pattern when they complete the questionnaire in different conditions. This behavioural preference can occur when the individual is going through change or transition, where there is uncertainty in their work or home life, where they have completed the questionnaire in a rushed manner or where they have changed their focus during completion on the context they are thinking about themselves in when answering the questions.

# Statement Selection

We use the strengths and order of the colours to select statements from the database applicable to each participant. Within the relevant database categories, the selection of statements is random to ensure each report is unique.

Within each category there are a certain number of preferred statements that include the participants name in order to make the report more personal.

Reports can be prepared from a selection of available categories e.g. Strengths, with either statements assembled in paragraphs (e.g. the Overview) or bullet points (e.g. Effective Communications), depending on the sections.

The most important validity for the report is the face validity; how well an individual participant identifies with their report. We are constantly assessing face validity to ensure the reports get the high percentage resonance we aim for. Ratings are typically over 85%.

In our initial testing our aim was for consistent resonance of over 80%. We tested the report output with thousands of individuals. This included a large cohort from Visa – a client with whom we had worked for many years. We asked them to rank C-me not only in terms of face validity, but also in comparison to other leading profiling systems. C-me was only taken out to market more widely when 90% of clients with whom we tested it said that their results were more accurate, and crucially more impactful, than peer comparisons [of 52 organisations who tested early outputs, 50 chose to use it as their behavioural insight tool of choice with no commercial incentive to do so]. At the point of launch resonance readings were consistently over 85%. Our current feedback is higher and we are continually seeking to hone statements.