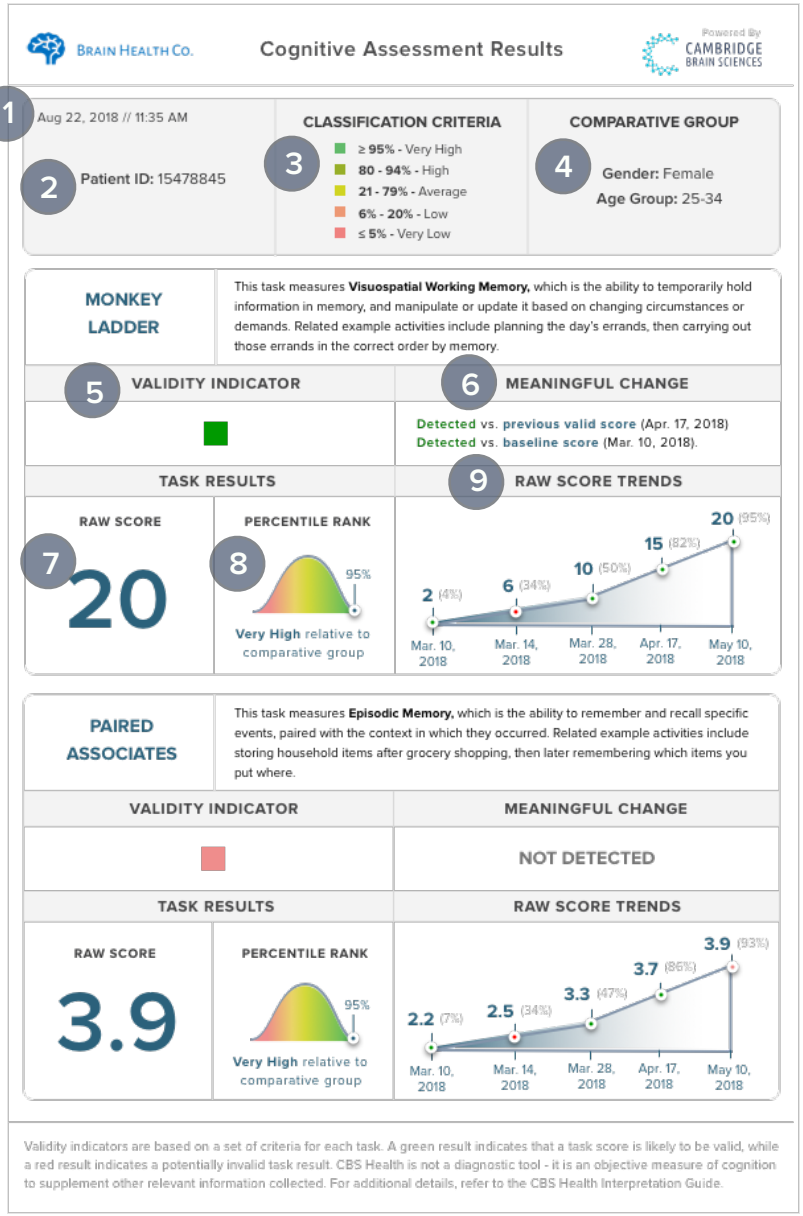
A. Core Components of the CBS Health Report
B. Raw Score Calculations
C. Task Validity Overview
D. Meaningful Change Overview
E. Answers to Common Interpretation Questions

Important Information Before You Read This Document

CBS Health provides a scientifically-validated and objective measure of a patient's cognition, however, it is not a diagnostic tool. CBS Health should be used in conjunction with other information and clinical judgement to reach conclusions regarding a patient's health. Ultimately, CBS Health does not replace the judgement of a practitioner and Cambridge Brain Sciences does not assume responsibility for the outcome of decisions made based on CBS Health data.

B. Core Components of the CBS Health Report

The CBS Health report is designed to be easy-to-read and shareable with patients. A brief overview of the components of the CBS Health report is shown below.



1 - The date and time the assessment was completed.

2 - The unique identifier given to the patient by the practitioner.

3 - Guidelines for how the patient's percentile rank compares to others in the comparative group.

4 - The subset of individuals within the Cambridge Brain Sciences normative database to whom the patient is being compared.

5 - Indicates whether task performance is within reasonable bounds using a set of validity parameters. If task results meet the criteria for validity, a green box will appear in this section. If task results do not meet the criteria for validity, a red box will appear instead. Validity criteria are outlined in Section C.

6 - Indicates whether changes in task results over time are potentially meaningful relative to a patient's previous and baseline assessment results. Conditions for meaningful change are outlined in Section D.

7 - The patient's raw score on the task. Section C outlines how raw scores for each task are calculated.

8 - The percentage of people in the patient's comparative group who attained the same result or lower on that task.

9 - A graphical presentation of raw score trends over time, including assessment completion dates and validity markers (red fill for potentially invalid data points, green fill for potentially valid data points). The trends graph will show a maximum of five of the most recent assessment results.

C. Raw Score Calculations

Raw scores are calculated differently for every task and these scores are not adjusted for age or gender. To learn more about the Cambridge Brain Sciences suite of tasks, refer to the Scientific Overview document.

Untimed Tasks (Digit Span, Monkey Ladder, Paired Associates, Spatial Span, Token Search)

For untimed tasks, raw score is calculated as the sum of the number of items in all successfully completed problems, divided by the number of successfully completed problems. For example, consider a Digit Span Task attempt in which a patient progresses through the following problems: 4 digits, 5 digits, [wrong answer], 4 digits, 5 digits, 6 digits, [wrong answer], [wrong answer]. The raw score will be the sum of all correctly remembered digits (4 + 5 + 4 + 5 + 6 = 24) divided by the total number of correct problems solved (5), resulting in a raw score of 4.8. This methodology provides greater granularity and a truer representation of the patient's performance compared to other more basic scoring methods, such as taking the highest level achieved on the task.

Timed Tasks Without Difficulty Level Adjustments (Double Trouble, Grammatical Reasoning)

For timed tasks in which difficulty does not adjust based on performance, the raw score is simply the number of correct answers minus the number of incorrect answers. For example, a score of 35 on Double Trouble may equate to 40 correctly answered problems and 5 incorrect problems.

Timed Task With Difficulty Level Adjustments (Feature Match, Spatial Planning, Odd One Out, Polygons, Rotations)

Feature Match and Rotations: the raw score is the number of points for correct answers, minus points subtracted for incorrect answers. Points correspond to the number of symbols; that is, problems with more symbols are worth more points than problems with fewer symbols. For example, a correctly answered problem with 6 symbols adds 6 points to the raw score, whereas an incorrectly answered problem with 7 symbols removes 7 points from the raw score.

Spatial Planning: each problem has a minimum number of moves needed to solve it. Points for a problem are equal to twice the minimum number of moves needed, minus the number of moves actually used. The raw score is the total points collected for all solved problems. A problem will end if patients make twice the number of moves needed, but points are not subtracted for failed problems. For example, completing a problem in 5 moves, when the minimum number of moves required is 3, results in a score of one for that problem: (2 x 3 minimum moves required) - 5 moves completed = 1.

Odd One Out: the raw score is the number of points for correct answers, minus points subtracted for incorrect answers.

Polygons: the raw score is the number of points for correct answers, minus points subtracted for incorrect answers. More difficult problems, which have more subtle differences between the polygons, are worth more points.

C. Task Validity Overview

CBS Health reports include a Validity Indicator—a simple and easy-to-read marker that conveys whether performance on a task is within reasonable bounds, helping you gain confidence that task scores reflect a patient's true cognitive state.

Conditions for Validity

Each task has a set of parameters* that must be met for the score to be considered likely valid. Utilizing the vast Cambridge Brain Sciences normative database, validity conditions have been computed for each task, and are shown below for reference:

Feature	Monkey Ladder	Spatial Span	Token Search	Paired Associates	Rotations	Polygons	Odd One Out	Spatial Planning	Grammatical Reasoning	Digit Span	Feature Match	Double Trouble**
Number of attempts	> 0	> 0	> 0	> 0	>= 7 and <= 36	> 0	>= 12 and <= 36	> 0	> 0	> 0	>= 14 and <= 37	> 0
Number correct	>= 11 and <= 105	>= 1 and <= 7	>= 1 and <= 12	>= 2 and <= 8	>= 5 and <= 30	>= 7 and <= 31	>= 8 and <= 34	>= 0 and <= 19	>= 6 and <= 46	>= 1 and <= 11	>= 13 and <= 33	>= 11 and <= 105
Number of errors	= 3	= 3	= 3	= 3	>= 0 and <= 12	>= 0 and <= 14	>= 0 and <= 15	>= 0 and <= 4	>= 0 and <= 9	= 3	>= 0 and <= 8	>= 0 and <= 37
Duration (seconds)	>= 60 and <= 277	>= 39 and <= 180	>= 24 and <= 751.3	>= 6 and <= 283	>= 89.5 and <= 90.5	>= 89.5 and <= 90.5	>= 179.5 and <= 180.5	>= 179.5 and <= 180.5	>= 89.5 and <= 90.5	>= 40 and <= 362	>= 89.5 and <= 90.5	>= 89.5 and <= 90.5
Max score	>= 5 and <= 11	>= 3 and <= 9	>= 2 and <= 14	>= 2 and <= 8	— — —	— — —	— — —	— — —	— — —	>= 3 and <= 12	— — —	— — —
Average Score	>= 3.4 and <= 7	>= 2.5 and <= 6.71	>= 2 and <= 9.08	>= 2 and <= 5	— — —	— — —	— — —	— — —	— — —	>= 3 and <= 8.4	— — —	— — —
Final Score	— — —	— — —	— — —	— — —	>= -4 and <= 294	>= -4 and <= 137	>= -2 and <= 33	>= 0 and <= 127	>= -1 and <= 45	— — —	>= 30 and <= 256	>= -5 and <= 102
Correct Score***	— — —	— — —	— — —	— — —	>= 20 and <= 326	— — —	— — —	— — —	— — —	— — —	>= 50 and <= 288	— — —
Max Level****	— — —	— — —	— — —	— — —	>= 5 and <= 18	— — —	>= 9 and <= 20	— — —	— — —	— — —	>= 6 and <= 16	— — —

*99% of scores in the Cambridge Brain Sciences normative database fall within the bounds of these parameters.

**Double Trouble includes additional measures of validity related to performance on different types of problems, such as 1) Congruent / Congruent (CC) problems, 2) Congruent / Incongruent (CI) problems, 3) Incongruent / Congruent problems (IC) and 4) Incongruent / Incongruent (II) problems. Additional validity metrics for the Double Trouble Task include:

- % of CC problems answered correctly: >= 60% and <= 100%
- Average reaction time in seconds when attempting CC problems: >= .752 and <= 4.81
- Average reaction time in seconds when attempting CI problems: >= .778 and <= 6.15
- Average reaction time in seconds when attempting IC problems: >= .761 and <= 5.65
- Average reaction time in seconds when attempting II problems: >= .784 and <= 6.01

***Correct score refers to the sum of all points obtained from successfully completed problems.

****Max level refers to the difficulty level of the hardest problem successfully completed.

D. Meaningful Change Overview

CBS Health reports also include a Meaningful Change indicator, which objectively determines whether the current assessment result represents a potentially significant change relative to two time-points: the patient's previous assessment results and the patient's baseline (e.g., initial) assessment results. Conditions for meaningful change for each of these cases are shown below. Note that Meaningful Change will only be calculated if a task score is likely valid according to the Validity Indicator.

1) **Conditions for Meaningful Change Relative to Baseline Assessment Results**

The top row of this chart indicates the attempt number, and the +/- indicates the required increase or decrease in the raw score for it to be considered potentially meaningful. For instance, if it is the patient's 6th time completing Polygons, they will need to see a raw score increase of 41.83 relative to their baseline (e.g., initial) assessment result in order for this change to be considered potentially meaningful.

Task	2nd		3rd		4th		5th		6th		7th		8th		9th		10th	
	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+
Monkey Ladder	-0.77	0.77	-0.69	0.90	-0.65	0.91	-0.70	1.00	-0.57	0.92	-0.55	1.00	-0.54	0.97	-0.57	1.13	-0.44	1.07
Spatial Span	-0.84	0.84	-0.81	1.04	-0.65	1.01	-0.75	1.12	-0.62	1.12	-0.84	0.84	-0.49	1.14	-0.96	0.96	-0.62	1.24
Token Search	-1.44	2.14	-1.45	2.26	-1.54	2.22	-1.23	2.27	-1.21	2.45	-0.96	2.35	-1.13	2.76	-0.99	2.78	-0.81	2.60
Paired Associates	-0.76	0.76	-0.64	0.93	-0.70	0.92	-0.64	0.87	-0.63	0.98	-0.63	0.95	-0.55	0.92	-0.78	0.78	-0.71	0.71
Rotations	-49.58	49.58	-45.70	62.34	-32.97	71.80	-31.37	70.18	-30.55	80.19	-25.48	83.30	-30.24	87.76	-33.23	94.94	-29.07	84.34
Polygons	-32.71	32.71	-26.85	37.40	-25.96	38.46	-27.42	42.96	-23.07	41.83	-26.27	44.58	-26.47	49.15	-25.23	53.86	-39.10	39.10
Odd One Out	-6.55	6.55	-6.41	6.41	-6.42	6.42	-5.87	5.87	-5.16	7.29	-4.86	8.09	-4.16	7.50	-4.20	6.93	-6.22	6.22
Spatial Planning	-10.45	19.94	-11.11	24.45	-6.96	25.54	-6.52	29.95	-4.02	31.65	-5.07	32.57	-5.42	34.90	-3.64	36.59	-5.99	37.33
Grammatical Reasoning	-6.58	6.58	-4.70	8.68	-5.13	9.59	-4.96	10.60	-4.90	10.18	-4.61	9.77	-4.33	10.20	-2.39	11.66	-3.88	12.19
Digit Span	-1.24	1.24	-0.90	1.29	-0.91	1.17	-0.79	1.32	-0.74	1.27	-0.82	1.44	-0.76	1.53	-0.74	1.39	-0.77	1.45
Feature Match	-47.44	47.44	-34.23	49.98	-32.73	52.30	-31.81	55.83	-30.48	57.25	-28.51	59.55	-29.79	66.09	-32.31	58.58	-24.44	54.01
Double Trouble	-12.25	22.48	-9.84	26.27	-9.20	29.54	-6.91	33.06	-3.50	33.99	-3.89	35.33	-6.71	39.85	-5.54	40.79	-2.81	45.10

2) **Conditions for Meaningful Change Relative to Previous Assessment Results**

The top row of this chart indicates which attempts are being compared, and the +/- indicates the required increase or decrease in the raw score for the change to be considered potentially meaningful. For instance, if it is the patient's 7th time completing Spatial Span, a raw score increase of 0.75 relative to their previous (aka, 6th) attempt will be considered potentially meaningful.

Task	1st to 2nd		2nd to 3rd		3rd to 4th		4th to 5th		5th to 6th		6th to 7th		7th to 8th		8th to 9th		9th to 10th	
	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+
Monkey Ladder	-0.77	0.77	-0.76	0.76	-0.68	0.68	-0.75	0.75	-0.78	0.78	-0.63	0.63	-0.79	0.79	-0.79	0.79	-0.62	0.62
Spatial Span	-0.84	0.84	-0.84	0.84	-0.88	0.88	-0.86	0.86	-0.81	0.81	-0.75	0.75	-0.91	0.91	-0.69	0.69	-0.95	0.95
Token Search	-1.44	2.14	-1.47	1.47	-1.48	1.48	-1.31	1.31	-1.23	1.23	-1.21	1.21	-1.18	1.18	-1.38	1.38	-1.03	1.03
Paired Associates	-0.76	0.76	-0.73	0.73	-0.74	0.74	-0.72	0.72	-0.69	0.69	-0.63	0.63	-0.74	0.74	-0.64	0.64	-0.75	0.75
Rotations	-49.58	49.58	-52.09	52.09	-48.17	48.17	-55.22	55.22	-46.38	46.38	-48.31	48.31	-45.05	45.05	-58.69	58.69	-49.69	49.69
Polygons	-32.71	32.71	-32.05	32.05	-32.39	32.39	-30.03	30.03	-29.63	29.63	-30.97	30.97	-34.10	34.10	-33.09	33.09	-31.21	31.21
Odd One Out	-6.55	6.55	-5.99	5.99	-5.59	5.59	-5.46	5.46	-5.36	5.36	-6.37	6.37	-5.36	5.36	-4.55	4.55	-4.47	4.47
Spatial Planning	-10.45	19.94	-14.90	14.90	-13.97	13.97	-15.42	15.42	-14.83	14.83	-14.89	14.89	-15.07	15.07	-13.22	13.22	-16.20	16.20
Grammatical Reasoning	-6.58	6.58	-6.42	6.42	-6.40	6.40	-6.36	6.36	-6.75	6.75	-6.94	6.94	-6.28	6.28	-5.79	5.79	-6.09	6.09
Digit Span	-1.24	1.24	-1.02	1.02	-1.07	1.07	-0.94	0.94	-0.96	0.96	-1.04	1.04	-1.01	1.01	-0.87	0.87	-0.93	0.93
Feature Match	-47.44	47.44	-45.85	45.85	-40.91	40.91	-45.46	45.46	-42.78	42.78	-43.26	43.26	-41.41	41.41	-38.79	38.79	-43.19	43.19
Double Trouble	-12.25	22.48	-16.23	16.23	-14.33	14.33	-15.25	15.25	-13.48	13.48	-14.87	14.87	-13.29	13.29	-12.40	12.40	-14.79	14.79

In the two charts above, statistical thresholds used are based on a two-tailed analysis, $p < 0.2$.

How are the conditions for meaningful change calculated?

The Meaningful Change Indicator compares the difference in an individual's performance on any given task, between two time points, to the variability in repeated measurements that would occur in the *absence* of meaningful change. The latter is estimated from a sample of healthy control subjects. The reliable change index uses the test-retest reliability and the standard deviation of scores (measured in a control sample) of a task to describe the range of possible differences that occur in repeat task completions. If an individual's change in performance from one time point to another is much larger than what is expected due to chance, then one can conclude that there was meaningful change.

Assessing meaningful change requires that these data have been obtained in a control sample. Cambridge Brain Sciences has a database of over 8 million test scores, and our normative database consists of 75,000 individuals. Most of these individuals have completed every task more than once, and more than 5000 people have completed every task more than 10 times. The interval between self-administered repeated assessments ranges from less than a day, to more than a month. This massive database allows us to characterize in the general population how performance on every task fluctuates naturally across a range of intervals. We are therefore able to quantify the bounds of what constitutes a meaningful change for every task, which are referenced in section D1 and D2.

E. Common Interpretation Questions

If a patient's percentile rank is "Low" or "Very Low" relative to their comparative group, is there cause for concern?

Firstly, ensure that "Low" or "Very Low" task results are marked as likely being valid within the Validity Indicator section of the CBS Health report. In certain rare instances, the patient may have misunderstood the task instructions or may have been distracted during the assessment.

In cases in which the Validity Indicator displays a likely valid task result, but scores are still within the "Low" or "Very Low" categories, you may wish to discuss with the patient the conditions under which the assessment was completed, and then choose to reassess that patient to obtain additional data. If the task results are still consistently to"Low" or "Very Low", further investigation may be warranted.

Ultimately, it is up to the practitioner to evaluate all available sources of information to decide the appropriate course of action. This may include continued treatment of the patient within the practitioner's area of expertise, reassessing the patient, or a referral to a neuropsychologist or neurologist.

Is it normal for a patient to score normally on most tasks, but very poorly on one or two tasks?

Unusually low scores on a few tasks may simply be due to chance or other irrelevant factors. In these cases, you may choose to re-assess the patient to gather additional data. However, each task is designed to measure distinct cognitive abilities, and tasks within one cognitive domain (e.g., memory) are largely independent from tasks within other cognitive domains (e.g., verbal ability). Therefore, it is normal for patients to exhibit strengths in one area of cognition, but weaknesses in another. In some cases, very low scores on a handful of tasks may indicate genuine weaknesses in one cognitive domain, and the areas of the brain associated with them. Examine patterns in the scores (e.g., was every low score a memory task?), and examine other information you may have collected to determine the appropriate next steps.

Is it possible for patients to get low scores even if their true capabilities are normal or high?

Obtaining a surprisingly low score is possible, and can arise due to a number of factors, such as patients being distracted, not understanding the task instructions, malingering (purposefully choosing wrong answers), or being unable to physically complete the assessment. In these rare occurrences, CBS Health may indicate that the task result was likely invalid.

Is it possible for patients to get high scores even if their true capabilities are normal or low?

Obtaining an artificially high score is highly unlikely, and only probable in cases where a patient is attempting to manipulate the results (e.g., using a pencil and paper to write down a sequence of digits in Digit Span). Taking the tests in this manner is detectable due to unusual patterns in the data, so in these cases, CBS Health may indicate that the task result was likely invalid.

Why do assessment results sometimes fluctuate so much from one session to the next?

Variability in task results is completely normal, given that the stability and reliability of neuropsychological tasks are inherently less than perfect. A measure of an individual's cognition will change upon repeated task completions as result of practice and novelty effects, regression to the mean, normal variation in performance, other random sources of error, and perhaps, meaningful change (e.g., deterioration resulting from a clinical condition). This is why we have incorporated a meaningful change indicator into CBS Health. A positive indication suggests that the change in the patient's score is larger than what we would expect to happen due to chance.

If a meaningful change is not indicated, does that mean cognition has stayed the same?

Not necessarily. The statistical analysis for meaningful change indicates when a variation in a task result is unlikely to be due to chance. It does *not* follow that a lack of significance means there is not meaningful change. Practically, this means that CBS Health reports are best interpreted as indicating when a change in a task result is unusual. If a result is not unusual, it is not definitive evidence that everything is normal.

My report is showing no meaningful changes detected, even though a condition for meaningful change has been met - why is this?

This is because the current task result is likely invalid. In order for there to be a meaningful change detected and displayed on a CBS Health report, the current task result must be shown as likely being valid.

How are the age ranges used for population comparisons determined?

Age ranges were chosen based on two objectives - i) minimizing significant differences in average performance within age ranges and ii) ensuring a large sample size data for that age range (and gender) to accurately compare the patient's score to the database.

How were the *classification criteria ranges determined?

The classification criteria were derived from standard statistical methodologies (i.e., standard deviations from the mean, using a normal distribution model), and transformed into percentile ranks to enhance usability and readability for practitioners and their patients, while still maintaining the same degree of scientific validity.

*See #3 in section B above