**Receiver operating characteristic curve**

Receiver operating characteristic (ROC) was used to assess the performance of binary support vector machine (SVM) classifier employed to classify the “yes” and “no” answer of the patient to respective questions. To draw a ROC curve for a patient, a contingency table was created for each session and also for the average of all the sessions, as shown in S3 Table.

Based on the results shown in S3 Table, true positive rate (TPR) and false positive rate (FPR) was calculated for each session and then for the average of all the sessions. The TPR and FPR was then used to draw ROC curve as shown in S2 Fig, S3 Fig, S4 Fig and S5 Fig.

Here we show the contingency table formed using the average of all sessions, separately for each patient, along with the chi-square statistics. While constructing the contingency table using the average of all sessions, training sessions and feedback sessions were averaged separately. The authors are aware of the limitations associated with the application of chi-square statistics to the present data, but still we performed the chi-square statistic just on exploratory basis. The contingency tables S4 Table, S5 Table, S6 Table, S7 Table, S8 Table, S9 Table, S10 Table and S11 Table provide the following information: the observed cell totals, (the expected cell totals) and [the chi-square statistic for each cell]. The value in the parenthesis in each cell shows the expected value for each cell calculated using the formula (*row total \* column total)* ***/****n*, where *n* is the total number of observations included in the table. The value in the square bracket in each cell shows the chi-square statistics value.

For S4 Table, the chi-square statistic is 150.84. The p-value is < .00001. The result is significant at p < .05

For S5 Table, the chi-square statistic is 22.42. The *p*-value is < .00001. The result is significant at *p* < .05.

For S6 Table, The chi-square statistic is 117.37. The p-value is < .00001. The result is significant at p < .05.

For S7 Table, The chi-square statistic is 16.21. The p-value is .000057. The result is significant at p < .05.

For S8 Table, The chi-square statistic is 108.62. The p-value is < .00001. The result is significant at p < .05.

For S9 Table, The chi-square statistic is 16.88. The p-value is .00004. The result is significant at p < .05.

For S10 Table, The chi-square statistic is 56.49. The p-value is < .00001. The result is significant at p < .05.

For S11 Table, The chi-square statistic is 10.31. The p-value is .001321. The result is significant at p < .05.

**Semantic concordance rate (SCR)**

Semantic paired sentences, such as “Paris is the capital of Germany” and “Paris is the capital of France”, and their classification were counted and analyzed against concordance between answers (in the above example, paired concordant answers are “false” and “true” respectively). The answering concordance of semantically paired sentences was estimated as the percentage of concordant answers over pairs' repetition.

The following analysis was carried out for each patient separately. For each sentences’ pair a contingency table was derived, then, True Positive Rate and False Positive Rate were computed (as described in the paragraph *Receiver operating characteristic curve*). Additionally, the True Negative Rate (TNR) was computed as follows:

 (1)

Averaged TPR and TNR across all sentences’ pairs were computed (i.e. *avgTPR* and *avgTNR* respectively), then the Semantic Concordance Rate (SCR) was defined as:

 (2)

To ascertain the consistency of the SCR estimates we applied formula to each single sentence’s pair (taking its TPR and TNR values into account) and checked whether median value across all pairs was different than 50% (with a random classifier it is plausible to expect a 50% of concordance across all sentences’ pairs). The Wilcoxon signed rank test was used because non-normal SCR distributions were detected with a chi-square goodness-of-fit test. S12 Table provides SCR estimates for each patient, including the percentage of sentences checked for semantic concordance (over the total amount of presented sentences).

Summarizing, the answering concordance between semantically paired questions was 68±1% (mean ± standard deviation), while its range was 67 - 78%.

Given the observed *fNIRS* classification accuracy (around 70%), the significant SCR estimates prove high concordance between the answers of semantically paired sentences covering a relevant subsample of the total presented sentences (more than 77% was checked against concordance).