Adults Can Be Trained to Acquire Synesthetic Experiences - Supplementary Information

Daniel Bor[1-2]*, Nicolas Rothen[1,3]*, David J. Schwartzman[1-2], Stephanie Clayton[1], and Anil K. Seth[1-2]

1. Sackler Centre for Consciousness Science, University of Sussex, Brighton, BN1 9QJ, UK.

2. Department of Informatics, University of Sussex, Brighton, BN1 9QJ, UK.

3. Department of Psychology, University of Sussex, Brighton, BN1 9QJ UK.

* These authors contributed equally to the work.

Correspondence:

Dr. Daniel Bor

Sackler Centre for Consciousness Science,

Department of Informatics,

University of Sussex,

Brighton,

BN1 9QJ,

UK.

+44 1273 678101

danielxbor@gmail.com

Methods

Test Tasks

Table s1 – Schedule of tests given at various stages of training

Pre-training Tests	Intermediate-Training Tests	Post-Training Tests	3 Month Follow up Tests
	Phenomenological Interview	Phenomenological Interview	Phenomenological Interview
Color Naming Stroop	Color Naming Stroop	Color Naming Stroop	Color Naming Stroop
Color Consistency Test		Color Consistency Test	
		Synaesthetic Conditioning	
		Synaesthetic Stroop	
Cattell Culture Fair IQ		Cattell Culture Fair IQ	

Training Tasks

For details of which tasks were administered at what training stage, see Table s3. All stimuli in the training tasks (except for the reading tasks) were shown at approximately 1.4 degrees viewing angle (3 cm high, 50 cm distance). Unless otherwise stated, all stimuli are presented on a light beige background. The same 13 letters were used in training for all participants, with the other 13 letters never associated with colors, effectively becoming control letters. The colors associated with each of the 13 trained letters are shown in Table s2. All tasks become progressively more difficult as performance improved, either on a trial-by-trial, or on a session-by-session basis.

Table s2 – List of training letters and their associated colors

Letter	Color	RGB Value
b	blue	0, 0, 153
d	brown	153, 51, 0
e	light green	102, 255, 102
g	green	0, 128, 0
i	light blue	0, 204, 255
o	orange	255, 161, 23
p	pink	255, 0 255
q	purple	128, 0, 128
r	red	255, 0, 0
u	grey	128, 128, 128
w	white	255, 255, 255
x	dark grey	85, 85, 85
y	yellow	255, 255, 0
Screen background	light beige	196,188,150

Passive Drill

Each letter was presented in its associated color for 1 s, followed by the entire background showing the associated color alone for 500 ms. This was repeated between 3 to 10 times, the number of repetitions randomly assigned for each letter. All 13 letters to be trained were presented in this cycle, in alphabetical order. The entire sequence of all letters was repeated twice. Participants were instructed to watch the letters carefully, and to try to memorize the presented letter-color associations. The entire block lasted approximately four minutes.

Early Delayed Match to Sample (E-DMTS)

For each trial, a random letter was assigned. The letter was initially shown in its associated color for 1.5s. Following a 3s interval, two colored squares were presented to left and right of the center, separated by 1cm, only one of which matched the color that the letter had just been presented in. Participants had up to 4s to respond by touching the square on the touch screen monitor that matched the color just presented. Feedback was given after each trial to indicate whether it was performed correctly. If the trial was performed correctly, the subsequent trial for that letter would include a foil square that was 20% more similar in color (in terms of RGB values) than before. If not, the subsequent trial for that letter would include a foil square that was 20% less similar in color. 30 trials were shown for each day's session. At the end of the session an average index of the final difficulty levels for all letters was shown to the subject, and was used at the start of of the next session for this task.

Late Delayed Match to Sample (L-DMTS)

This was identical to E-DMTS, except the letters were all shown in black, rather than in their associated colors. Subjects therefore had to recall the associated colour and then match it, making the task more challenging.

Early Span

For each trial, a sequence of random letters was presented. Each letter was on screen for 500 ms, with a 250 ms gap between letters. Letters were presented in their associated color. No letter appeared more than once in any given sequence. 500 ms following the final letter in the sequence, the word "respond" was displayed, and immediately after this, 13 colored squares were horizontally arranged across the middle of the screen, in random order, representing the 13 letter-color associations. Participants attempted to touch the squares on

4

the touch screen monitor in the same order as the colors of the letter sequence just presented. Feedback was given after each trial to indicate whether it was performed correctly or not. If the trial was performed correctly, the subsequent trial would be longer by one letter, otherwise it would be shorter by one letter, with a minimum sequence of 1 letter possible. 20 trials were given per session. At the end of the session an average of the session's trial sequence lengths was shown to the subject, which was used to set the beginning of the next session for this task, rounded down to the nearest integer.

Late Span

This was identical to the early span, except that random colored squares instead of letters were initially presented in sequence, involving a subset of the 13 possible colors that the trained letters were associated with. The response stage involved black letters presented in alphabetical order. Participants attempted to touch the letters in the appropriate sequence, which required converting the previous sequence of colored squares into a letter sequence, according to the trained color-letter associations.

Advanced Span

This task was identical to the early span task, except that black letters were used instead of colored letters. Colored squares were still the targets, as in the early span task.

Colors to Letters Speed Test

For each trial, a colored square was centrally presented, one third from the top of the touch-screen monitor. One third from the bottom of the screen the 13 training letters were presented in alphabetical order. Participants were required to touch as fast as possible the letter associated to the presented color with. The response triggered the next trial, with a different color presented, 250 ms later. At the end of the session, feedback was given on the average speed for correct trials. 100 trials were given per session.

Letters to Colors Speed Test

This was identical to the Colors to Letters Speed Test, except that on the top a single letter was presented instead of a single color square, and on the bottom the 13 possible trained colors were presented in random order for each trial.

Reading Easy

For each training session a hard copy of a 1-2 page A4 general interest article was given to subjects to read. On the first session, a single letter of the 13 to be trained was printed in its associated color. For each subsequent session an additional letter was colored, until on the 13th training session all 13 letters were colored, the untrained letters always being shown in black.

Reading Advanced

After the 13th training session instead of substituting black training letters with colored ones, a single letter was replaced by a corresponding colored square after each training session until all letters were replaced by colored squares on the 26th training session e.g. A Mann■■ M■ss■■n t■ Ma■s.

Synesthesia N Back

For each trial, either a random black letter or colored square was centrally presented (random 50% chance of each) for 1 second, with a 2 second gap between trials. On stimulus presentation, participants were required to press the space bar only if the color or letter on the current trial matched the color or letter on the trial n back. Note that color square trials could be matched with letter trials, if the color associated with the letter was the same as the color square. 50% of trials, randomly assigned, were match trials. 50 trials were presented in a given session. The first session was 1 back. At the end of the session percentage correct

feedback was given to the participants. If participants scored over 90% on a session, then the n would be increased by 1 for the next session. For occasional sessions for some subjects, due to misunderstanding by a junior research assistant, the improvement criterion was set at 80% instead.

Backward Span

For each trial, a sequence of random letters was presented. Each letter was on screen for 500ms, with a 250ms gap between letters. Letters were presented in their associated color. No letter appeared more than once in any given sequence. 500ms following the final letter in the sequence, the word "respond" was shown on the screen, and immediately after this, 13 black letters were horizontally arranged across the middle of the screen, in alphabetical order. Participants attempted to touch the squares on the touch screen monitor in the opposite order to the sequence just presented. Feedback was given after each trial to indicate whether it was performed correctly or not. If the trial was performed correctly, the subsequent trial would be longer by one letter, otherwise it would be shorter by one letter, with a minimum sequence of 1 letter possible. 20 trials were given per session. At the end of the session an average of the session's trial sequence lengths was shown to the subject, and this level of difficulty was then set for the beginning of the next day's session on this task, rounded down to the nearest integer.

Spelling Task

For each trial, a horizontal sequence of color squares was presented a third from the top of the screen for 500 ms. This sequence, if translated to the appropriate letters, would spell a word. After a 250ms gap, the 13 training letters were presented a third from the bottom of the screen and participants were required to touch the letters in the correct order, to spell the word. Feedback on whether the word was correct or wrong

was given after every trial. 20 trials were presented per session. At the end of the session participants were given feedback about the overall percentage accuracy achieved. There were 4 difficulty levels (level 1 = 2 or 3 letter words, level 2 = 4 letter words, level 3 = 5 letter words and level 4 = 6 or 7 letter words), with level 1 given on the first session and a given difficulty level kept constant throughout the session. If participants scored above 90% on a given difficulty level, then they would be given the next more difficult level during the next session. For occasional sessions for some subjects, due to misunderstanding by a junior research assistant, the improvement criterion was set at 80% instead.

Book homework

During the pre-test subjects were asked to choose a book that they would like to read in their own time over the course of the study from Project Guttenberg (http://www.gutenberg.org/). Each book was converted into a standard text file and the letters colored according to the 13 letter-color associations used in the study. The file was saved as a .pdf and emailed to subjects. Subjects were encouraged to read the book at home throughout the study and were assessed on completion by being asked to give a description of the plot of the book.

Table s3 – Schedule of training tasks given throughout the training stage

Task	Trial # Per Training Session	Week 1	Week 2	Week 3	Week 4	Week 5	Week 6	Week 7	Week 8	Week 9
Passive Drill	20	✓								
Early Delayed Match to Sample	30	✓	✓							
Late Delayed match to sample	30			✓	✓					
Early Span	20	✓	✓	✓						
Late Span	20	✓	✓	✓	✓	✓				
Reading Easy	----	✓	✓	✓	✓	✓				
Letters to Colors Speed Test	100		✓	✓	✓	✓	✓			
Colors to Letters Speed Test	100		✓	✓	✓	✓	✓			
Advanced Span	20				✓	✓	✓	✓	✓	✓
N-Back	50					✓	✓	✓	✓	✓
Backward Span	20						✓	✓	✓	✓
Reading Advanced	----						✓	✓	✓	✓
Spelling task	20							✓	✓	✓

Results

Training effects

Training gain on the Early and Late Delayed Match To Sample tasks was measured by the change in average session color difference between the probe and foil, as determined by the subjects' staircase performance. Training gain on the Early, Late, Advanced and Backward Span Tasks was measured by the change in average span length presented per session, again determined by the subject's staircase performance. Training gain on the N back and Spelling tasks was measured by the change in difficulty level (n for N Back and difficulty level reflecting length of words to spell in the spelling task) between sessions, again dependent on a between session staircase method. Finally, training gain on the Colors to Letters and Letters to Colors Speed Tests was measured by the difference in average reaction time per session. See table s3 and figure s1 for details of linear regression functions and the trajectory of improvement on each task based on the above criteria.

Table s4 – Linear regression statistics for all training tasks

TASK	Linear r^2	F	df	p	Effect size (Cohen's f^2)
Early Delayed Match to Sample	.774	408.508	1, 119	<0.001	3.42
Late Delayed Match to Sample	.105	16.126	1, 137	<0.001	0.12
Early Span Task	.239	59.451	1, 189	<0.001	0.31
Late Span Task	.467	272.944	1, 312	<0.001	0.876
Advanced Span Task	.061	26.859	1, 413	<0.001	0.06
Backward Span Test	.083	25.914	1, 288	<0.001	0.09
N Back	.681	716.723	1, 335	<0.001	2.13
Letters to Colors Speed Test	.102	34.795	1, 307	<0.001	0.11
Colors to Letters Speed Test	.092	31.536	1, 310	<0.001	0.10
Spelling Task	.361	116.415	1, 206	<0.001	0.56

Figure s1 – Training performance across sessions (±S.E.), averaged over all participants.
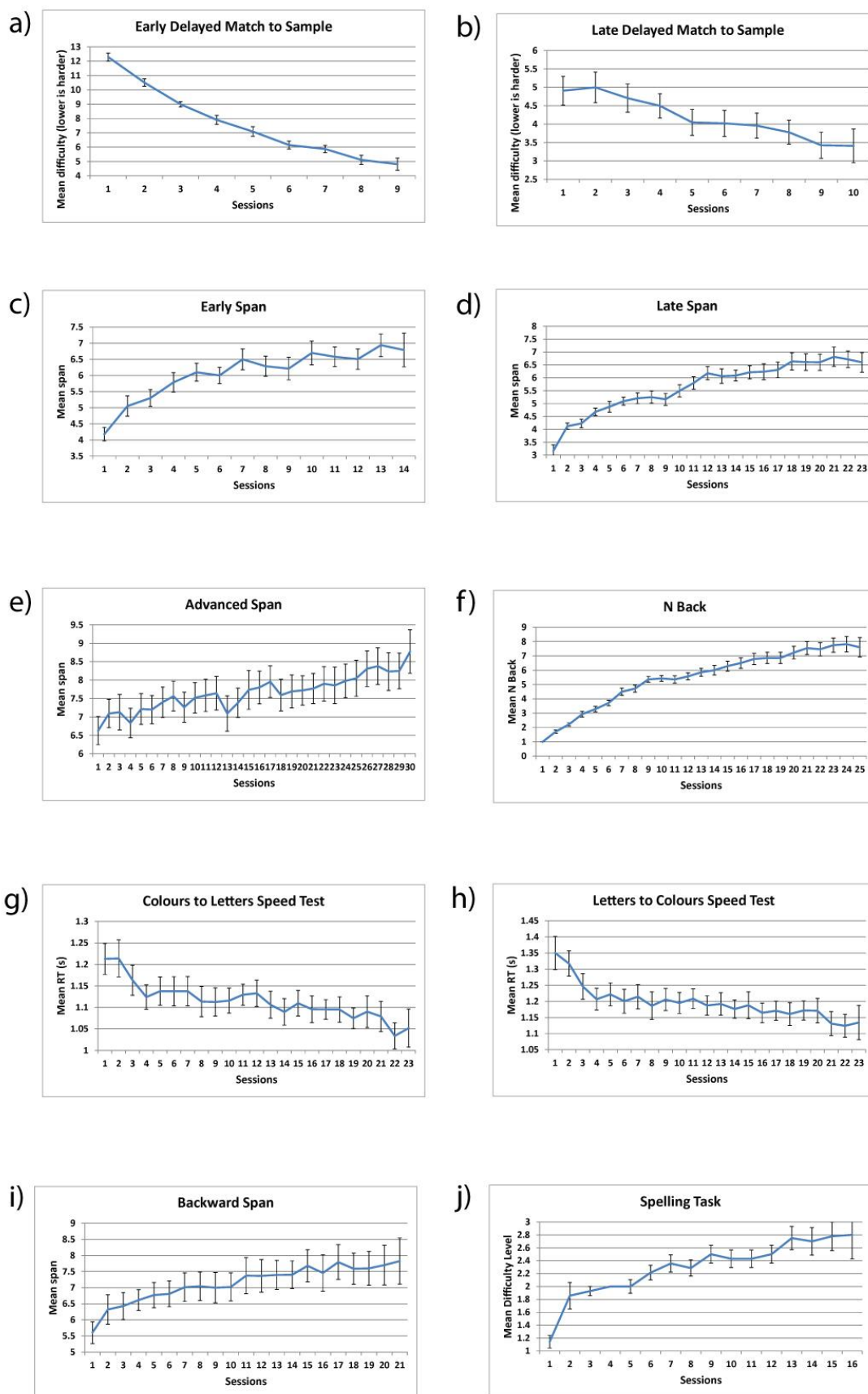
Tests

Intermediate Stage

Table s5 – Summary of the phenomenological interview at the intermediate stage of training

Subject No.	Synesthetic Experiences	Letter personas	Example of experiences
1	No	Yes, doesn't like d,x,u	
2	Yes	Yes, i-"peaceful", x-"aggressive"	"I see it [colors] on signs in my mind... If I think of a word in my mind the letters appear colored... If I think about the spelling of a word or imagine a word in my mind the letters are colored" Subject noticed these experiences 2-3 times a day
3	No	Yes, darker colors "meaner"	
4	No	Yes, w-"neutral", x-"evil"	
5	Yes	No	"Walking through Pevensey [a set of buildings on campus] I saw the letter R on a poster and thought of the color red" Had this experience >10 times, mainly on signs around campus.
6	Yes	Yes, y-"happy", u-"annoying", i-"easy going", d-"upsetting", w-"shocking"	"In the advanced span [training task] as soon as the letter disappears I experience the color in my mind"
7	No	Yes, e,i-"nice", q-"mild", r-"not nice", x-"horrible"	
8	Yes	Yes, b,y,x-"happy colors"	"In my mind's eye I see the color as the appropriate letter"
9	Yes	Yes, e-"don't like this letter", x-"not a nice letter"	"I see red in my head when I see the letter R"
10	No	Yes, d-"the beginning", r-"power", w-"nirvana"	
11	Yes	Yes, x-"friendly"	"When thinking about which letters to use in the tasks I see the color filling my mind"
12	Yes	No	"When reading a sign on campus I saw all the letter E's color green on the sign"

13	No	Yes, r-reminds the subject of his cousin	
14	Yes	Yes, E-eccentric, U-subject feels pity for the letter	"I am starting to have the experience of the color associated with the letter in the external world" Subject reported that these experiences were very frequently and occurred more often when she was reading.

Post Training

Self-Reports of Imagery Ability

Three synesthesia experts independently rated the extent of synesthetic phenomenology, based on the participant's post-training reports, as a percentage of genuine synaesthesia. These percentages were then converted to a rank score for each participant per ranking expert to minimize individual biases in ranking, and the mean rank between the three experts for each participant was finally used as an index of synesthetic phenomenology. This was then compared with VVIQ and SUIS scores to test for correlations between imagery self-report and phenomenology. No significant correlation was found either with VVIQ (r^2 = 0.04, F (1,12) = 0.52, p>0.1, effect size: Cohen's f^2 = 0.04) or SUIS (r^2 = 0.05, F (1,12) = 0.68, p>0.1, effect size: Cohen's f^2 = 0.05). No significant correlations were either found between the Color Naming Stroop Task and VVIQ (r^2 = 0.19, F (1,12) = 2.77, p>0.1, effect size: Cohen's f^2 = 0.23) or SUIS (r^2 = 0.02, F (1,12) = 0.20, p>0.1, effect size: Cohen's f^2 = 0.02), or between the Synesthetic Stroop Task and VVIQ (r^2 = 0.01, F (1,12) = 0.07, p>0.1, effect size: Cohen's f^2 = 0.01) or SUIS (r^2 = 0.01, F (1,12) = 0.12, p>0.1, effect size: Cohen's f^2 = 0.01).

Phenomenological Intensity And Training And Test Results.

In addition to self-reports of imagery, the extent of synesthetic phenomenology, as rated by three experts, was compared with other training and test results. No significant correlations were found with any of the training tasks (based on percentage improvement between the first and last sessions). The correlations of phenomenology against either the Color Naming Stroop and Synesthetic Stroop Task scores was also not significant. However, the correlation between phenomenology ranking and color consistency scores approached significance (r^2 = 0.27, F (1,12) = 4.53, p>0.055, effect size: Cohen's f^2 = 0.37).

Cattell Culture *IQ Test*

Figure s2 – Cattell Culture Fair IQ Test for Trained and Control Groups.

a)
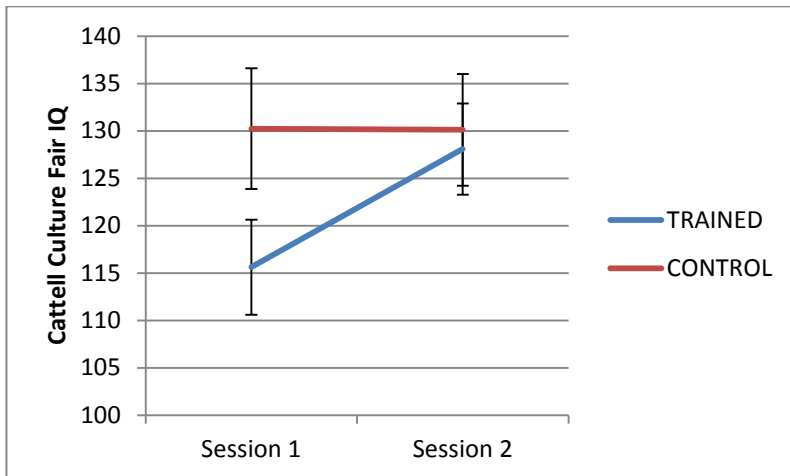
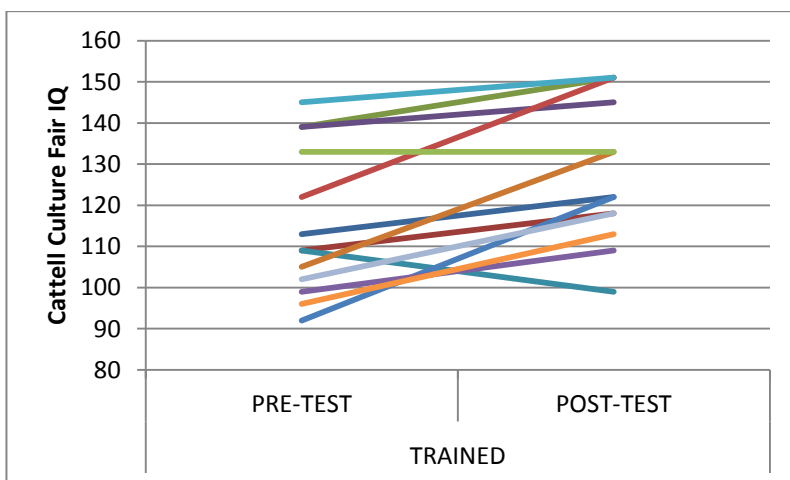b)

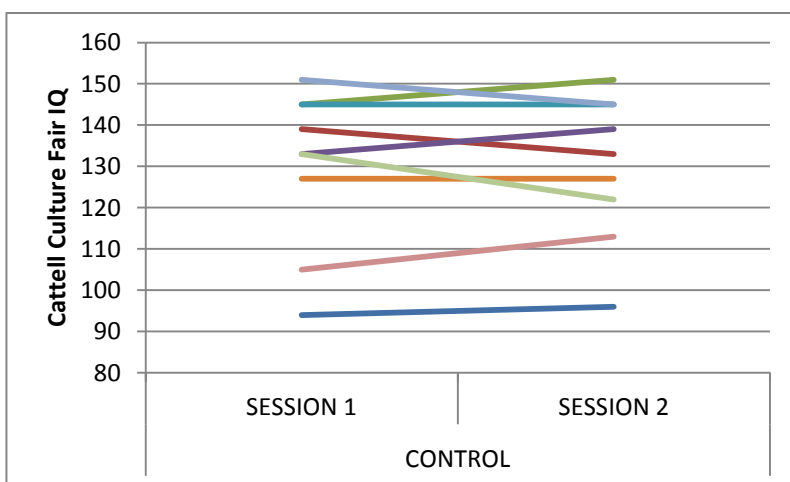
c)

Figure s2 – a) Group IQ comparison (±S.E.) between synesthetically trained participants and controls. b)

Individual subject gain scores for trained participants and c) controls

Color Consistency Test

Examining the single subject effects, rather than group effects, following training 11/14 participants were below the (CIELUV) synesthetic threshold for trained letters and 4/14 were below for untrained letters, as shown in figure s3.

Figure s3 – Color consistency scores on the Color Consistency Test following training for individual subjects
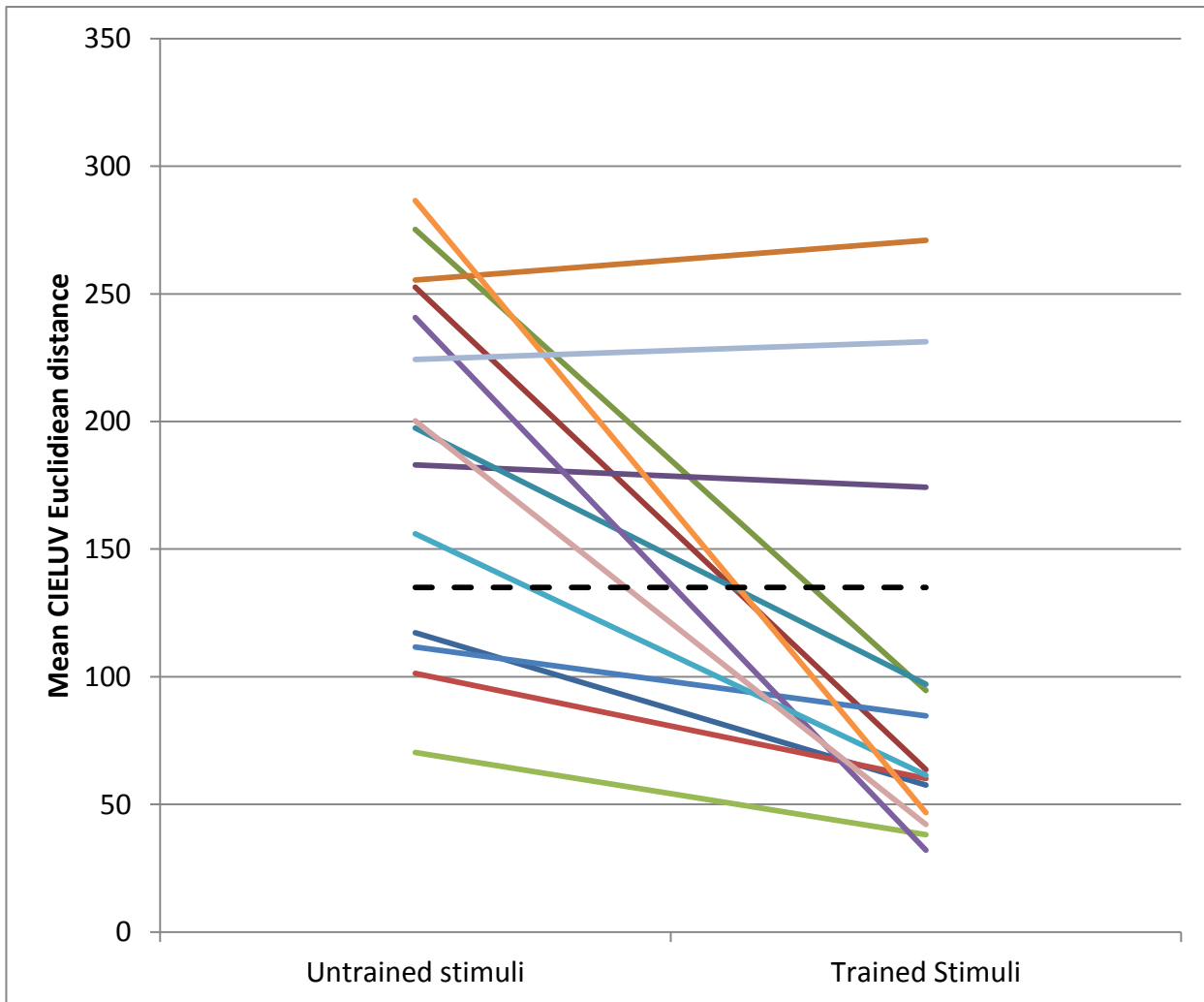
Figure caption: Color consistency scores (±S.E.), based on the CIELUV Euclidian distance algorithm, using the online Color Consistency Test for the 13 trained and 13 untrained letters after training only. Each subject is represented by a single solid line. A lower score reflects increased color consistency. Values below the dashed line are normally assumed to signify genuine synesthesia.

The color consistency test was additionally analysed, split according to whether the associations were semantic (7 letters) or non-semantic (6 letters) in nature. This was done in case the semantic associations were inflating the consistency results. There was a main effect of session ($F(1,11)=20.78$ $p<0.001$, effect size: partial $eta^2 = 0.654$), no main effect of trial type (semantic versus non semantic to-be-trained/trained letters) ($F(1,11)=0.492$ $p>0.1$ effect size: partial $eta^2 = 0.044$), and no significant interaction ($F(1,11)=0.702$ $p>0.1$ effect size: partial $eta^2 = 0.014$). Therefore, there was no differential consistency effect for the semantic, compared to the non-semantic color-letter associations.

Figure s4 – Color consistency scores on the Color Consistency Test, split by whether the associations were semantic or not
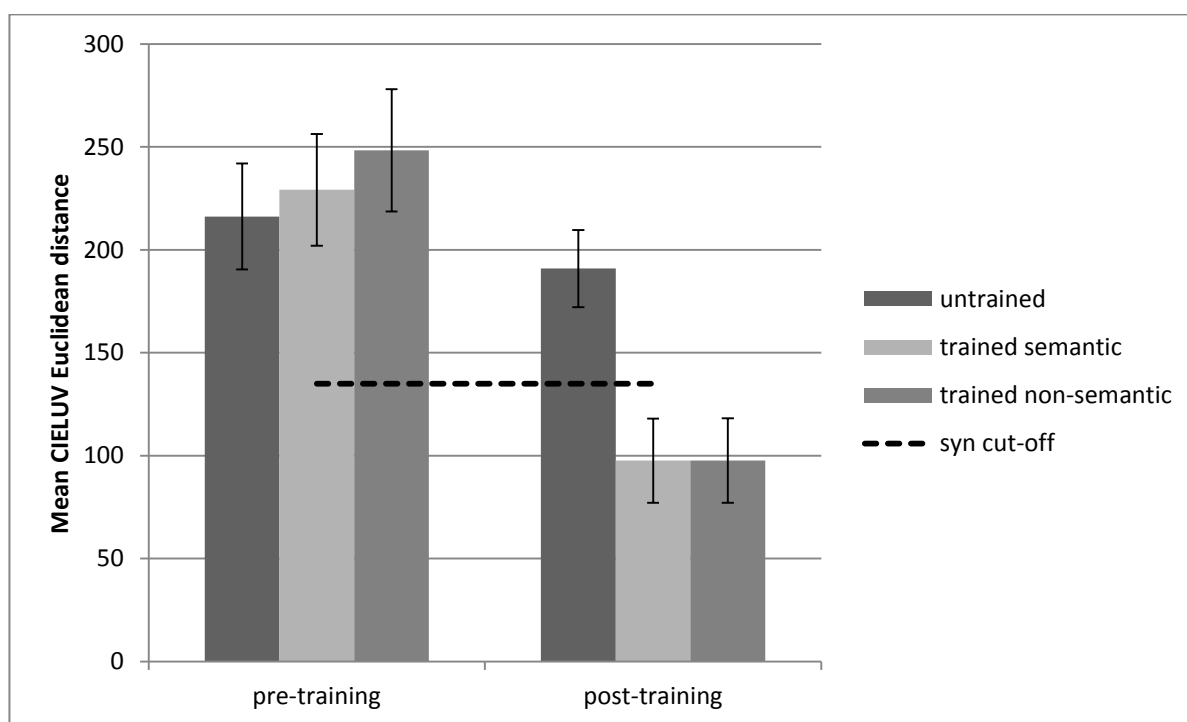
Figure caption: Color consistency scores (±S.E.), based on the CIELUV Euclidian distance algorithm, using the online Color Consistency Test for the 13 trained (split by semantic association or not) and 13 untrained letters before and after training. A lower score reflects increased color consistency. Values below the dashed line are normally assumed to signify genuine synesthesia.

Finally, in order to test whether our results are inflated due to those standard and potentially easily remembered colors, the color consistency test was analysed according to whether the associations were standard colors (i.e. all RGB values either 0 or 255; 4 letters) or non-standard colors (9 letters). There was a main effect of session ($F(1,11)=21.92$ $p<0.001$, effect size: partial $eta^2 = 0.67$), no main effect of trial type (standard versus non-standard colors) ($F(1,11)=0.601$ $p>0.1$ effect size: partial $eta^2 = 0.052$), and no significant interaction ($F(1,11)=1.03$ $p>0.1$ effect size: partial $eta^2 = 0.086$). Therefore there was no inflation in consistency scores due to the standard colors.

Figure s5– Color consistency scores on the Color Consistency Test, split by whether the colors for associations were standard or non-standard
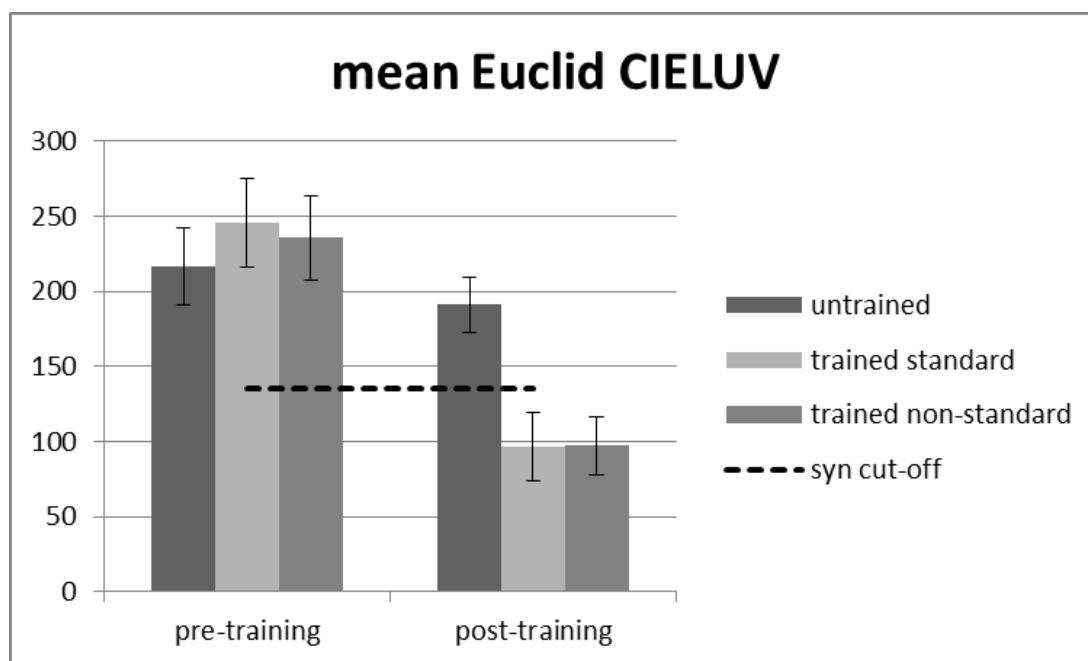
Figure caption: Color consistency scores (±S.E.), based on the CIELUV Euclidian distance algorithm, using the online Color Consistency Test for the 13 trained (split by standard/non-standard color associations) and 13 untrained letters before and after training. A lower score reflects increased color consistency. Values below the dashed line are normally assumed to signify genuine synesthesia.

3 Month Follow up

Color Naming Stroop

Note that two participants were unable to attend the 3 month follow up session. There was still a significant congruency effect (t(11)=7.58, p<0.001, effect size: Cohen's d = 0.84), which was marginally smaller than the post-training congruency effect, but not significantly so overall (post-training mean congruency effect: 106ms; follow up mean congruency effect: 96ms; t(24)=0.48, p>0.1, effect size: Cohen's d = 0.19) or specifically for the semantic component (post-training mean semantic congruency effect: 190ms; follow up mean congruency effect: 174ms; t(24)=0.50, p>0.1, effect size: Cohen's d = 0.20).

Figure s6 – 3 Month follow up Color Naming Stroop Results (±S.E.)