Consciousness and Cognition 33 (2015) 375–385

Contents lists available at ScienceDirect

Consciousness and Cognition

journal homepage: www.elsevier.com/locate/concog

Validating a standardised test battery for synesthesia: Does the Synesthesia Battery reliably detect synesthesia?

D.A. Carmichael [a,b,c,e,*], M.P. Down [b], R.C. Shillcock [a,b], D.M. Eagleman [d], J. Simner [a,e]

[a] Dept. of Psychology, University of Edinburgh, 7 George Square, EH8 9JZ, UK
[b] Institute for Adaptive & Neural Computation, University of Edinburgh, 10 Crichton Street, EH8 9AB, UK
[c] Division of Psychiatry, University of Edinburgh, Royal Edinburgh Hospital, EH10 5HF, UK
[d] Department of Neuroscience, Baylor College of Medicine, Houston, TX, USA
[e] School of Psychology, University of Sussex, Falmer BN1 9QH, UK

ARTICLE INFO

Article history:
Received 6 March 2014

Keywords:
Synesthesia
Validation
Test
Assessment
Consistency
Prevalence
Grapheme-colour

ABSTRACT

Synesthesia is a neurological condition that gives rise to unusual secondary sensations (e.g., reading letters might trigger the experience of colour). Testing the consistency of these sensations over long time intervals is the behavioural gold standard assessment for detecting synesthesia (e.g., Simner, Mulvenna et al., 2006). In 2007 however, Eagleman and colleagues presented an online 'Synesthesia Battery' of tests aimed at identifying synesthesia by assessing consistency but within a single test session. This battery has been widely used but has never been previously validated against conventional long-term retesting, and with a randomly recruited sample from the general population. We recruited 2847 participants to complete The Synesthesia Battery and found the prevalence of grapheme-colour synesthesia in the general population to be 1.2%. This prevalence was in line with previous conventional prevalence estimates based on conventional long-term testing (e.g., Simner, Mulvenna et al., 2006). This reproduction of similar prevalence rates suggests that the Synesthesia Battery is indeed a valid methodology for assessing synesthesia.

1. Introduction

Synesthesia is an inherited condition in which everyday stimuli trigger unusual secondary sensations. For example, synesthetes listening to music might see colours in addition to hearing sound (Ward, Huckstep, & Tsakanikos, 2006). One particularly well-studied variant is *grapheme-colour synesthesia*, in which synesthetes experience colours when reading, hearing or thinking about letters and/or digits (e.g., Simner, Glover, & Mowat, 2006). Despite being first reported over two hundred years ago (by Sachs, 1812; see Jewanski, Day, & Ward, 2009) synesthesia was initially an under-researched and poorly-understood area of human experience until the last decades of the 20th century. A significant factor in the elevation of synesthesia as a tractable topic was the realisation – and subsequent empirical confirmation – that synesthetes' experiences could be verified behaviourally by the fact that they remain conspicuously stable over time (Baron-Cohen, Wyke, & Binnie, 1987; Jordan, 1917). Specifically, synesthetes tend to be highly consistent when reporting their synesthetic sensations for any given stimulus. For example, if the letter J triggers the colour pale blue for a given synesthete, she will tend to repeat that J is pale blue (not green, not yellow, etc.) when repeatedly tested over days, months and even years. Indeed,

* Corresponding author at: 3.34 Informatics Forum, University of Edinburgh, 10 Crichton St, Edinburgh EH8 9AB, UK.
E-mail address: d.a.carmichael@ed.ac.uk (D.A. Carmichael).

http://dx.doi.org/10.1016/j.concog.2015.02.001
1053-8100/

one study was able to show that synesthetic sensations had remained consistent over at least three decades (Simner & Logie, 2008).

This stability of responses over time is considered one of the central features of synesthesia and is routinely verified in almost every publication on the subject (e.g., Asher, Aitken, Farooqi, Kurmani, & Baron-Cohen, 2006; Baron-Cohen, Burt, Smith-Laittan, Harrison, & Bolton, 1996; Rich, Bradshaw, & Mattingley, 2005; Ward & Simner, 2003; but see Simner, 2012). In other words, while a wide range of behavioural approaches have been employed to assess the nature of the synesthetic experience, experimental methodologies aiming to *validate* synesthesia have almost exclusively focussed on the feature of consistency. Hence, researchers selecting synesthete participants for study first verify the genuineness of each case by requiring their synesthetes to demonstrate high levels of consistency over time compared to non-synesthete controls (e.g., Asher et al., 2006; Baron-Cohen et al., 1996; Simner et al., 2006). Controls are tested on analogous associations (i.e., they invent colours for the 26 letters, say, and then attempt to recall these colour associations later) and typically perform significantly worse than synesthetes.

Although more than a hundred contemporary studies rely on this test of consistency for genuineness, the particular instantiation of the test has varied widely. For example, a wide range of methods have been used to elicit synesthetic colours: participants have indicated these by either giving verbal descriptions (e.g., Ward, Simner, & Auyeung, 2005), written descriptions (Simner, Glover et al., 2006), using Pantone© swatch colour charts (Asher et al., 2006), electronic colour charts (Simner, Harrold, Creed, Monro, & Foulkes, 2009) or even computerised colour pickers offering extensive palettes of >16 million colours (e.g., Simner & Ludwig, 2012). In this way, synesthesia research has used varying methods, which in turn might raise difficulties for researchers when trying to meaningfully compare data.

Despite this superficial variability however, the test of genuineness has nonetheless tended to rely on one key shared feature: synesthetes must outperform controls over fairly lengthy re-test intervals. Consider, for example, the most widely cited large-scale screening for synesthesia (Simner, Mulvenna et al., 2006) in which a large sample of participants were opportunistically recruited from the communities of Edinburgh and Glasgow Universities, and individually assessed for synesthesia. Participants first indicated by questionnaire whether they believed they experienced synesthesia, and those who reported in the affirmative were asked to provide their synesthetic associations (e.g., the colours of letters). These participants were then retraced after considerable time had passed (on average 6.0 months) and were asked in a surprise retest to re-state their associations. A group of controls without synesthesia performed an analogous task but were re-tested after only two weeks. Synesthetes were able to significantly out-perform controls even though much time had passed and the deck was effectively stacked against them. Methodologies such as this allow confident detection of genuine synesthetes because the surprise retest over lengthy intervals places the performance of synesthetes beyond the usual abilities of the average person. The drawback to this methodology, however, is that the task is extremely time-intensive to perform, and risks a high drop-out rate if synesthetes become untraceable at retest.

Perhaps for this reason, one of the most important developments in the methodology of synesthesia validation came with the introduction in 2007 of an alternative version of the test of genuineness. Eagleman and colleagues produced the *Synesthesia Battery*, a toolbox of online tests which provides a standardised set of questions, tests and quantitative scores to assess a range of synesthesias (Eagleman, Kagan, Nelson, Sagaram, & Sarma, 2007). This battery is again based on internal consistency in that synesthetes are validated by high consistency within their own synesthetic associations, stated repeatedly. However, consistency is measured within a single test session lasting only approximately 10 min. Specifically, synesthetes log on to the testing site (www.synesthete.org) and specify which form(s) of synesthesia they experience. The testing platform then presents their triggering stimuli (e.g., the 26 letters) one by one in randomized order, and participants are required to select their synesthetic colour for each trigger. Each stimulus (e.g., letter) is presented three times each, and a score is generated to quantify the consistency of participant's responses (e.g., did the participant choose the same/similar colours each of the three times she saw a particular letter?) This score represents the geometric distance in RGB (red, green, blue) colour space, where R, G, and B values are all normalised to lie between 0 and 1. If the mean overall score of colour-distance is less than 1, the participant is classified as a synesthete; if the score is 1 or higher, the degree of inconsistency classifies the participant as a non-synesthete. However, it remains an open question whether this limited retest interval is sufficient to truly distinguish synesthetes from non-synesthetes.

In the current study we assessed the validity of the Synesthesia Battery by using it to test almost 3000 randomly sampled subjects for grapheme-colour synesthesia. Our aim was to establish the prevalence of grapheme-colour synesthesia by this method. This will allow us to evaluate the Synesthesia Battery by comparing this prevalence – obtained by assessments within in a single test session – to the most widely accepted previous estimate of the prevalence of grapheme-colour synesthesia based on the standard *longitudinal* test–retest method (Simner, Mulvenna et al., 2006). If the Synesthesia Battery is just as effective a method for detecting synesthesia as the more standard long-term retest method, we anticipate an equivalent prevalence of grapheme-colour synesthesia across both methods. In carrying out our study, we chose to evaluate grapheme-colour synesthesia in particular for several reasons: it is one of the most common forms of synesthesia (Simner, Mulvenna et al., 2006), it is particularly well-understood in behavioural terms, it lends itself readily to online testing, and those who experience it typically demonstrate the high levels of consistency expected from synesthetes (compared to other variants, whose more complex concurrents may make them more difficult to assess via consistency alone; see Simner, Gäartner, & Taylor, 2011 for discussion). It was not our intention to change or try to improve upon the method made available by Eagleman et al. at www.synesthete.org. Rather, we attempted to simply replicate their test and methodology and then evaluate how it performs in comparison to a conventional longitudinal test–retest method.