



**HAL**  
open science

# The prevalence and diagnosis of ‘stereoblindness’ in adults less than 60 years of age: a best evidence synthesis

Adrien Chopin, Daphne Bavelier, Dennis Michael Levi

► **To cite this version:**

Adrien Chopin, Daphne Bavelier, Dennis Michael Levi. The prevalence and diagnosis of ‘stereoblindness’ in adults less than 60 years of age: a best evidence synthesis. *Ophthalmic and Physiological Optics*, 2019, 39, pp.66 - 85. 10.1111/opo.12607 . hal-03919204

**HAL Id: hal-03919204**

**<https://hal.science/hal-03919204>**

Submitted on 2 Jan 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



INVITED REVIEW

# The prevalence and diagnosis of 'stereoblindness' in adults less than 60 years of age: a best evidence synthesis

Adrien Chopin<sup>1,2</sup>, Daphne Bavelier<sup>1,2</sup> and Dennis Michael Levi<sup>3</sup> 

<sup>1</sup>Faculte de Psychologie et Sciences de l'Education, University of Geneva, <sup>2</sup>Campus Biotech, Geneva, Switzerland, and <sup>3</sup>School of Optometry, University of California, Berkeley, USA

**Citation information:** Chopin A, Bavelier D & Levi DM. The prevalence and diagnosis of 'stereoblindness' in adults less than 60 years of age: a best evidence synthesis. *Ophthalmic Physiol Opt* 2019; 39: 66–85. <https://doi.org/10.1111/opo.12607>

**Keywords:** amblyopia, stereoblindness, stereopsis, stereopsis tests, strabismus

*Correspondence:* Dennis Michael Levi  
E-mail address: [dlevi@berkeley.edu](mailto:dlevi@berkeley.edu)

Received: 26 October 2018; Accepted: 23 January 2019

## Abstract

*Purpose:* Stereoscopic vision (or stereopsis) is the ability to perceive depth from binocular disparity - the difference of viewpoints between the two eyes. Interestingly, there are large individual differences as to how well one can appreciate depth from such a cue. The total absence of stereoscopic vision, called 'stereoblindness', has been associated with negative behavioural outcomes such as poor distance estimation. Surprisingly, the prevalence of stereoblindness remains unclear, as it appears highly dependent on the way in which stereopsis is measured.

*Recent findings:* This review highlights the fact that stereopsis is not a unitary construct, but rather implies different systems. The optimal conditions for measuring these varieties of stereoscopic information processing are discussed given the goal of detecting stereoblindness, using either psychophysical or clinical stereotests. In that light, we then discuss the estimates of stereoblindness prevalence of past studies.

*Summary:* We identify four different approaches that all converge toward a prevalence of stereoblindness of 7% (median approach: 7%; unambiguous-stereoblindness-criteria approach: 7%; visual-defect-included approach: 7%; multiple-criteria approach: 7%). We note that these estimates were derived considering adults of age <60 years old. Older adults may have a higher prevalence. Finally, we make recommendations for a new ecological definition of stereoblindness and for efficient clinical methods for determining stereoblindness by adapting existing tools.

## Introduction

Stereoscopic vision (or stereopsis) is the ability to perceive depth from the difference of viewpoints between the two eyes. Stereopsis is functionally important in humans for distance estimation in realistic environments,<sup>1,2</sup> object recognition,<sup>3</sup> object prehension,<sup>4</sup> object placement,<sup>5,6</sup> or tasks involving fine hand movements.<sup>7,8</sup> Stereopsis has been said to impart a sense of volume ('the space between objects') to the visual world.<sup>9</sup> However, not everyone can appreciate depth from disparity. The total absence of stereoscopic vision is called 'stereoblindness' and a partial dysfunction is referred to as stereo-impairment, stereo-anomaly or stereo-deficiency. Quantifying stereoacuity and identifying stereoblindness is critical for evaluating the integrity of the visual pathways, for assessing the impact of

visual training and for treating patients with binocular anomalies.

## What causes stereoblindness?

Stereoblindness and stereo-impairment can arise from a number of distinct causes. For example, in their review of the effects of brain damage on stereopsis, Holmes & Horax<sup>10</sup> described a patient with parietal cortex damage who often walked into objects because his world appeared flat.<sup>11</sup> It has been long known that patients with strabismus and/or amblyopia frequently have reduced or absent stereopsis.<sup>12,13</sup> Moreover, stereopsis is frequently impaired in patients with age-related macular degeneration (AMD).<sup>5</sup> However, the proportion of the population that is reported to be stereoblind or stereo-impaired is highly dependent on

the way in which stereopsis is measured.<sup>14–17</sup> Therefore, in the first part of this review, we discuss the optimal conditions for measuring stereoacuity with the goal of detecting stereoblindness and stereo-impairment. In the second part, we will see how these variations in measuring stereoblindness influence the prevalence of the defect and we will attempt to extract the most accurate estimate. The current review is focused on adults <60 years of age; however, we acknowledge that the decline of stereoacuity with increasing age is a major issue even in healthy adults.<sup>18</sup> The case of adults 60 years of age and above would be important to consider on its own in future work given that stereoacuity declines with cataract and the progression of glaucoma,<sup>19</sup> and is improved by cataract surgery (particularly second eye cataract surgery).<sup>20,21</sup>

### How to measure stereoacuity for determining stereoblindness?

#### What are the challenges for measuring stereoblindness?

There is currently no standard definition and no simple diagnostic test for stereoblindness, thus patients are typically operationally classified as ‘stereoblind’ when they fail to respond correctly to the largest disparity in a given test (e.g. Randot circles or Butterfly stereotests). Simmerman<sup>22</sup> notes that ‘one of the most potentially frustrating aspects of stereopsis testing is finding a valid test whose results can be easily interpreted’. Determining stereoblindness is even more challenging than measuring stereoacuity. In this review, we focus on how to measure residual stereoacuity and determine stereoblindness, avoiding potential errors that may arise from the measure itself.

There are two kinds of potential errors when measuring stereopsis: to miss it when it is present (false negative) or to detect it when it is absent (false positive). To avoid both of these errors, a test needs to: (1) attribute a non-zero score if any residual stereopsis is present, and (2) avoid any possibility of deceiving the test, i.e., to exclude the possibility that other sources of information may allow correct responding in the absence of genuine stereopsis. The first condition is met by a test aimed at catching impaired but present stereopsis. While this appears to be a simple goal, in practice it is complicated by the existence of different sub-systems of stereovision (e.g. stereo from divergent or convergent disparities), which partially overlap. Thus, a requirement will be that the test uses a large range of testing conditions, so that it can measure separately the different sub-systems of stereovision. Another pitfall is to count any error as a test failure: it happens with tests requiring, for example, four correct responses out of four trials. Later, we refer to these tests as *too strict*. The second condition is met by employing a test that has a very low probability of

obtaining a non-zero score by guessing, as in any good psychometric tool. An added difficulty, however, is that there are many non-stereoscopic cues in stereoacuity tests. It is imperative that the extent to which these non-stereoscopic cues may help to pass the test be properly addressed. Below we discuss how to reach these goals.

#### How to measure residual stereoacuity optimally?

To increase the sensitivity of stereotests, Westheimer<sup>23</sup> recommends measuring stereoacuity with high target luminance (30–1000 cd.m<sup>-2</sup>) and the highest possible contrast, given stereoacuity is proportional to contrast.<sup>24,25</sup> Items should be spaced by the smallest separation<sup>26</sup> compatible with minimal crowding (around 10<sup>/27</sup>), have sharp edges (therefore, avoiding Gabor patterns), and be presented for at least 400 ms.<sup>25</sup> In addition, in normal vision, stereoacuity is optimal at the fovea, and with no pedestal disparity.<sup>26</sup> Normal observers have increased sensitivity to small convergent disparities<sup>28,29</sup> compared to large divergent ones. An added feature of stereopsis is that it improves quickly with practice.<sup>30,31</sup> This test calls for a rapid method, in order to avoid confounding learning effects. However, observers should also be provided with extended time to familiarise them with the test and instructions.

Participants who are classified as stereoblind typically suffer from other visual defects which may mask residual stereo-vision.<sup>32</sup> To avoid such masking, large items (around 25') should be used<sup>33</sup> and observers should have their vision corrected to their best acuity. We note that in amblyopic patients and those with AMD, stereopsis may be reduced or absent in foveal vision, but be spared in peripheral vision,<sup>34</sup> so it may be important to test either with a large field, or to test both centrally and in the periphery. Another source that confounds is matching errors due to inaccurate vergence. To limit those, the use of binocular fusion locks seems helpful.

#### Testing multiple stereoscopic sub-systems

Stereovision is not a unitary system: it is comprised of several partially independent sub-systems.<sup>35</sup> It is possible that only one sub-system is impaired for some observers. If only this sub-system is tested, a participant may be categorised as stereoblind when he/she is only stereo-impaired within a given sub-system. To avoid this scenario, we recommend testing more than one sub-system at the same time. Below, we review the different sub-systems.

##### *Convergent vs divergent disparity systems*

There is an ongoing debate between two views on the processing of convergent (or crossed/near) and divergent (or uncrossed/far) disparities. According to the first view, each

is processed in separate systems (see Mustillo<sup>36</sup>). According to the opposite view,<sup>37,38</sup> there is a continuum of disparity sensitivities.<sup>39,40</sup> Whether one view or the other is correct, the important fact is the following: some observers can show normal stereoacuity for one type of disparity and be stereoblind to the other.<sup>15,38,41–44</sup> As a result, either both types of disparity should be present in the stereotest, or each type should be measured separately. There is a common misconception that a test depicting a hovering target measures sensitivity to convergent disparities. However, the geometry of convergent and divergent disparities depends ultimately on the current state of convergence.<sup>45</sup> If the observer fixates on the target, its disparity is zero and the background has divergent disparities; yet, if the observer fixates on the background, the target now has convergent disparities while the background has zero disparity. Therefore, measuring convergent and divergent systems separately is only possible if vergence movements are controlled or if the stimuli are briefly presented (<200 ms). The clinical stereotests reviewed here all assess both divergent and convergent disparities, concurrently.

#### *Bright and dark stereopsis systems*

It has been argued that bright and dark polarities are processed independently by the stereopsis system.<sup>46,47</sup> The idea came from the observation that random-dot stereograms made of both bright and dark dots on a grey background are easier to process than those made of only bright or only dark dots.<sup>46</sup> This idea of independence was recently challenged by the demonstration of an artefact when using stereograms of bright and dark dots,<sup>48</sup> but the artefact was only present when dots were not allowed to overlap. We therefore recommend using mixed polarities with no item overlap. Interestingly, bright and dark polarities are tightly bound to the convergent and divergent systems. Indeed, observers who are blind to convergent disparities when using bright-on-dark targets become blind to divergent disparities when using dark-on-bright targets.<sup>49</sup> Thus, a test with both convergent and divergent bright-on-dark-targets will, by design, cover the cases of dark-on-bright stereoblindness. Given that every clinical stereotest in this review includes both disparities, they also test for both polarities. The issue is therefore irrelevant for these tests.

#### *Motion-in-depth vs static disparity systems*

The relationship between motion-in-depth and stationary stereopsis systems is complex. There is evidence in favour of independent systems, mostly (1) an absence of correlation between performance on clinical stereotests and motion-in-depth tests,<sup>50</sup> and (2) several cases of double dissociations, in which motion-in-depth scotomas exist with normal stereopsis<sup>51</sup> and residual motion-in-depth exists in the absence of stereopsis in strabismic observers.<sup>52,53</sup>

However, evidence of tight relationships between the two systems also exists, with (1) correlational evidence that motion-in-depth relies independently on two mechanisms, one based on motion and the other on disparities,<sup>54</sup> and (2) a study showing that strabismic patients with residual motion-in-depth are also the ones with residual stereopsis.<sup>55</sup> We recommend either testing only for stationary stereopsis, to avoid masking a potential stereoblindness with a residual motion-in-depth ability that would rely partially on interocular speed differences (motion system), or testing both subsystems separately.

#### *Absolute vs relative disparity systems*

Absolute disparity (the difference between the angle subtended by the target at the two entrance pupils of the eyes and the angle of convergence) is important for judging the depth distance of an object from one's self, and is the cue for convergence. Relative disparity is the difference between the absolute disparities of two objects. There is converging evidence that relative disparities are calculated from absolute disparities and that the two systems are therefore not independent.<sup>32,56,57</sup> However, conscious depth perception shows little or no direct access to absolute disparities (the absolute disparity anomaly<sup>57</sup>). Some brain-damaged patients can also show large impairments of stereopsis from relative disparities while still be able to normally (but poorly) access absolute disparities.<sup>58</sup> However, we recommend using a relative disparity task, given such a task is also an absolute disparity task and relative disparity sensitivity is better than absolute disparity sensitivity for the typical observer. Currently, all clinical tests measure both disparities concurrently.

#### *Coarse vs fine disparity systems*

Stereopsis is optimal with disparities that fall within Panum's area, in which stimuli can be fused (fine stereopsis), but stereopsis also exists with disparities so large that stimuli appear diplopic (coarse stereopsis<sup>59,60</sup>). Within Panum's area, and for some diplopic disparities, magnitude estimates are possible (patent stereopsis), while larger disparities only trigger signed depth (qualitative stereopsis<sup>61</sup>). There is accumulating evidence that fine and coarse stereopsis might be processed by separate stereo-systems (see Wilcox & Allison<sup>62</sup>) while others argue that there is a continuum of disparity sensitivities.<sup>37,38</sup> Coarse disparities develop earlier and are more robust to amblyopia than fine disparities.<sup>63,64</sup> There is also a double dissociation with observers able to see small, but not large, disparities and the converse.<sup>39,41,65</sup> However, the dissociations could also reflect a single disparity system with different ranges of disparity sensitivity for different individuals.

In all cases, to be sure to measure any residual disparity sensitivity, (1) the stereotest should offer a range of

disparities wide enough to include both fine and coarse disparities and, (2) all disparities should be administered regardless of the result for the other disparities. We arbitrarily define fine disparities as the ones within Panum's area and coarse disparities as the one outside of it. As a limit for Panum's area, in *Table S1*, we use the disparity that triggers the same number of 'single' and 'double' reports, and that is around 900".<sup>66</sup> Staircases should be avoided given that the coarse-fine dichotomy breaks the first assumption necessary to make the staircases converge, that proportions of correct responses are a monotonic function of stimulus level.<sup>67</sup> The coarse-to-fine method of limits suffers the same limitation. Indeed, stereo-sensitivity increases with increasing disparity up to some value, and then diminishes to zero. This maximum disparity compatible with stereopsis is known as  $D_{\max}$ , and it varies depending on stimulus conditions.<sup>68</sup> A disparity greater than  $D_{\max}$  will result in no depth perception even in observers with normal stereoacuity, which limits the range of large disparities that can be used to determine the presence or absence of stereopsis. None of the clinical tests reviewed here contains disparities exceeding  $D_{\max}$ .

#### *Transient vs sustained disparity systems*

Several studies have provided evidence for two systems with different spatio-temporal sensitivities,<sup>69</sup> a transient system that requires large disparities and flashed presentations,<sup>69,70</sup> and a sustained one that prefers small disparities presented for at least 1 s.<sup>71</sup> Others have argued that the dichotomy could simply reflect the fact that that depth information accumulates over time and that transient stimuli are therefore harder.<sup>38</sup> Indeed, no case of deficit with an impaired sustained system and an intact transient system has yet been reported.<sup>35</sup> However, the transient system can extract opposite-polarity disparities that cannot be processed by the sustained system,<sup>71</sup> demonstrating the dichotomy between the two systems. Ideally, stereoacuties should therefore be tested separately at short and long stimulus durations; however, that is generally not practical with current clinical methods.

#### *First vs second-order disparity processing*

There is evidence that grating stereograms are processed by two mechanisms with distinct spatial and temporal properties.<sup>68,72,73</sup> The dichotomy matches the coarse-fine and transient-sustained dichotomies.<sup>62</sup> More specifically, a first-order mechanism seems to process the fine disparity information from carrier signals and increases its precision over time. A second-order mechanism prefers the coarse disparity and transient envelope signal of stimuli.<sup>62</sup> Importantly, there is also a size-disparity correlation: in larger stimuli, stereo-depth can be perceived for larger disparities than in smaller stimuli.<sup>38</sup> Therefore, to recruit the first-

order system, small stimuli with fine disparities should be presented for long durations. To recruit the second-order system, large stimuli with coarse disparities should be presented for short durations.

#### *Global vs local stereopsis*

Global stereopsis occurs when a global solution to the binocular correspondence problem of several elements in depth is required, e.g. in random-dot stereograms.<sup>74</sup> Global stereopsis can trigger the perception of illusory contours. Local stereopsis refers to the matching of isolated items (e.g. lines or dots). While adaptation experiments favour a single mechanism for both situations,<sup>75</sup> lesion studies provide evidence either for one<sup>76</sup> or two separate mechanisms,<sup>77,78</sup> with no established double dissociation. Given that only impaired global stereopsis with intact local stereopsis was found, either local stereopsis or both should be tested. However, we will see later that it is very challenging to test local stereopsis while avoiding monocular cues.

#### *Near vs distance testing*

While stereoacuity is quite similar at distance and near in individuals with normal vision, it may vary substantially in individuals with strabismus, if the degree of misalignment of the two eyes varies with viewing distance.<sup>79</sup> Thus, ideally, testing would be done at both near and far distances.

#### *Instantaneous vs sequential stereopsis*

While *instantaneous stereopsis* is the estimation of relative disparities during steady fixation, there is a form of *sequential stereopsis*<sup>80</sup> that requires at least two different eye fixations on different depth planes. A specific version of sequential stereopsis, called delta-vergence,<sup>81</sup> allows the observer to extract relative disparities from changes of vergence. The larger the distance between two items, the more the visual system relies on delta-vergence than on instantaneous stereopsis.<sup>81</sup> Observers unable to use instantaneous stereopsis may still be able to use delta-vergence (and vice versa) because vergence can rely on absolute disparities while those disparities are inaccessible to other visual modules.<sup>57,82</sup> Therefore, the ideal stereoblindness test would allow delta-vergence thanks to long presentations and either large disparities or a stimulus spanning a large visual field.

#### **Chance levels**

Chance performance levels (i.e. the likelihood of responding correctly based purely on guessing) are often unreported or not considered in studies evaluating whether stereotests can be contaminated by non-stereoscopic cues (see next section for details on those cues). High chance levels make the test unlikely to be useful for detecting stereoblindness. If an observer can guess the correct response to

a test with one chance out of three, for example, then up to 33% of totally stereoblind observers could be misclassified as having stereopsis. Therefore, stereotests with chance levels as low as possible are preferred, ideally one-tenth of the expected prevalence for stereoblindness (we use 0.7% in *Table S1*). Likewise, the proportions of observers passing the test monocularly should always be compared to its chance level. In our example, if the test can be passed monocularly with a 33% probability, it does not demonstrate that monocular cues are present, but simply that participants can guess at chance level. However, if the test can be passed monocularly with a significantly higher probability than chance, then it indicates that monocular cues are present. An added complication is that chance levels to pass the tests at the lowest non-zero score can be non-intuitive to calculate. For example, chance levels are often reported as a probability of guessing at the starting disparity. When the test is not started at the largest disparity, a failure at that disparity usually leads to repeating the test at a larger disparity, which increases the chance level significantly. Chance levels are available in *Table S1*, along with the details of the calculations. Manufacturers sometimes do not give precise instructions so that each researcher or clinician uses a different criterion for passing the test. Changing the criterion often leads to large changes in chance level. For example, if there are four alternatives, with a criterion of four correct consecutive responses to pass the first level, chance level is 0.4%. If the criterion is three correct responses out of four attempts, chance level is now 5%.

### The challenge of non-stereoscopic cues

Stereopsis tests may contain cues to depth other than stereoscopic disparity cues: the *non-stereoscopic cues*. These cues first need to be removed from the stereotests before genuine stereoscopic depth performance can be measured. There are two families of non-stereoscopic cues: monocular cues and binocular cues, the latter having been only recently demonstrated.<sup>83</sup>

Monocular cues encompass any aspect of the stereotest allowing above chance performance and that would be perceptible with only one eye. These include perspective,<sup>84</sup> motion parallax,<sup>85</sup> accommodation,<sup>86</sup> or the alternating fixation described by Archer.<sup>87</sup> The best way to show the presence of monocular cues in a stereotest is to determine whether observers can pass the test monocularly with a higher probability than chance. In the review presented here, we do not consider studies that simply assumed that patients with serious binocular dysfunctions are totally stereoblind because such patients may have residual stereovision.<sup>88</sup>

Tests based on natural depth with real items are susceptible to the use of motion parallax (when moving the

head<sup>85</sup>), size change and accommodation. Monocular cues can often be removed from tests by using random-dot stereograms rather than edge stereograms. In edge stereograms, the disparity information coincides with the edges of the item in depth. In random-dot stereograms, patterns of random dots are given different binocular disparities depending on whether they are in the target area (a specific shape) or in the background. The limit between the target and the background is not an edge of the target or background item, but a disparity-defined illusory contour. However, random-dot stereograms introduce potential *binocular non-stereoscopic cues*. The possible use of binocular non-stereoscopic cues to pass stereotests was first mentioned by Cooper,<sup>89</sup> and Charman and Jennings.<sup>90</sup> Cooper<sup>89</sup> simply noted that binocular non-stereoscopic cues are probably present in random-dot stereograms: he coined them *decorrelation cues*. Charman and Jennings realised that some participants could identify shapes in random-dot stereograms when a stereotest was rotated 90°, making the disparities vertical. Vertical disparities usually do not produce depth but binocular images can then become incompatible and trigger binocular rivalry - a noticeable alternation of each eye's image, often in a piecemeal manner.<sup>91</sup> The authors noted that the rivalry in a rotated test could help participants to identify the shapes in depth even when the test is not rotated. The rivalrous area could even float on top of the fused background, a phenomenon called *rivaldepth*.<sup>92</sup> Furthermore, random-dot stereograms can produce binocular lustre,<sup>93</sup> or a diplopic or confusion area<sup>94</sup> that can behave as binocular non-stereoscopic cues. Charman and Jennings<sup>90</sup> did not provide any empirical evidence to substantiate their claim, but we recently demonstrated the existence of binocular non-stereoscopic cues in two different stereotests.<sup>83</sup> Most stereo-normal observers were able to use these binocular non-stereoscopic cues to deceive stereotests. The importance to be given to those cues is relative. While detecting binocular (de)correlation is conceptually different from detecting a signed estimate of disparity, the same neural mechanisms may mediate the two abilities. Thus, a task which can be performed by detecting binocular decorrelation (or other non-stereo binocular cues) may also tap into stereopsis more generally. This is suggested by the observation that stereoblind observers do not seem to combine left and right eye signals non-linearly, as would be required if they were sensitive to interocular correlation,<sup>95</sup> at least for large unstructured fields. On the other hand, a fraction of a small sample of stereoblind observers was able to use binocular non-stereoscopic cues to deceive stereotests<sup>83</sup>: 27% (3/11) for the Circle or Shape part of the Randot, 100% (2/2) for the Shape part of the Butterfly test.

In summary, non-stereoscopic cues encompass monocular cues and at least four kinds of binocular cues: binocular

lustre, diplopia/confusion, binocular rivalry and rivaldepth. All these cues could in theory be used to deceive stereotests in the total absence of stereoscopic vision (Table 1). Binocular non-stereoscopic cues are more likely to appear in tests with dense textures, whose binocular misalignments can produce rivalry, diplopia/confusion, lustre and rivaldepth. One way to render these cues useless is to ask whether the shape is protruding or recessing (sign estimate), or which of two items is more protruding or recessing (magnitude estimate). Therefore, we classified the presence of binocular non-stereoscopic cues for stereotests that have not been previously evaluated as follows in Table S1: likely when (1) dense textures are present, and (2) the task is a detection of depth or of an odd item, or a shape identification; unlikely when only (1) or (2) is true; and very unlikely if neither (1) nor (2) is true. Next we will review how clinical and laboratory stereotests deal with non-stereoscopic cues and whether clinical stereotests are appropriately designed to test for stereoblindness.

### Clinical stereotests

In clinical settings, stereoacuity is usually measured with commercial clinical stereotests,<sup>96,97</sup> often with poor results in diagnosing binocular dysfunctions.<sup>98–100</sup> A more detailed description of most of the following clinical tests can be found elsewhere.<sup>101–105</sup> The test-retest reliability of many stereotests is very poor, requiring a change of as much as a factor of four in order to exceed test-retest variability.<sup>106</sup> We now review the presence of non-stereoscopic cues for the most common tests along with their strengths and weaknesses (Tables S1 and S2). In this review, when we say that a test is passed, we mean that the observer responds correctly to the largest available disparity, except when clearly stated otherwise.

### Two-needle, Howard and Howard-Dolman tests

One solution to avoiding monocular cues is to make them unreliable, as when comparing the distance of two real needles seen from far away (6 m). This is the approach of the Howard test, Howard-Dolman test and their two-needle

**Table 1.** The relations between the different categories of cues to pass stereotests. Only the latter column (binocular stereoscopic cues) is of interest when measuring stereoacuity

| Non-stereoscopic cues        |                    | Stereoscopic cues  |
|------------------------------|--------------------|--------------------|
| Monocular cues               | Binocular cues     |                    |
| Motion parallax              | Binocular rivalry  |                    |
| Perspective                  | Diplopia/Confusion | Absolute disparity |
| Shadows                      | Binocular lustre   | Relative disparity |
| Alternating fixation (. . .) | Rivaldepth         | Delta-vergence     |

variants.<sup>2</sup> Such tests can produce excellent stereoacuties, on the order of a few arc-seconds. Unfortunately, the assumption that monocular cues are unreliable at distance is not correct: the presence of monocular cues always increases distance estimation acuity when added to the disparity cue, even when viewed from as far as 12 m.<sup>107</sup> While size change does not seem useful at that distance,<sup>108</sup> Howard<sup>2</sup> showed that the minimal performance on his test could be reached for 70% of his observers when tested with one eye only, demonstrating a probability of monocular cues far above chance level (binomial test with  $n = 10$ ,  $p = 0.003$ ). However, none of 12 observers<sup>109</sup> could pass the two-needle test when monocularly patched. Unfortunately, we could not test the overall difference with the chance level given their procedures resulted in very different chance levels. Indeed, Howard's original procedure suffers from a high chance level (25.4%; Table S2) limiting the interpretation in case of a success to the test. Binocular non-stereo cues have not been investigated but they are unlikely because no dense texture is used and it is a depth-ordering task. Finally, the test is impractical in most indoor environments with limited space and commercialisation has been discontinued.

### Titmus test

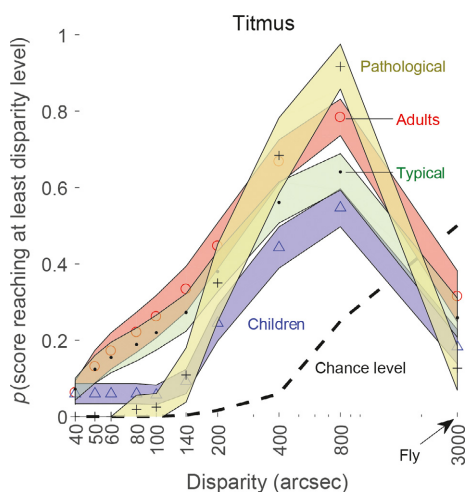
The Titmus stereotest<sup>110</sup> (<https://www.stereooptical.com/products/stereotests-color-tests/original-stereo-fly/>) uses polarised glasses with vectograph images to simulate depth. In this review, we will focus on the Fly and Circles portion of that test, ignoring the Animal portion.

The Titmus Fly is a vectographic image of a fly that presents the largest disparity available in clinical stereotests. In an influential study,<sup>111</sup> Simons and Reinecke claimed that monocular cues were present in these portions of the test. Indeed, large proportions of neurotypical and strabismic adult observers,<sup>112,113</sup> and of children with monofixation syndrome<sup>114</sup> reported the fly floating in depth, even though the fly was viewed monocularly and could therefore not appear in stereoscopic depth. However, none of the proportions differed significantly from chance. Although it is likely that monocular cues exist in the Fly, the main problem is the high chance level (50%). Recently, two studies<sup>115,116</sup> used a rotated viewing condition to improve testing with the Titmus Fly. In that condition, between 19% and 48% of their samples could pass the test, a result that could also be explained by the high chance level.

The Circles portion of the Titmus test (or Wirt circles) is among the most commonly used stereotests in the USA.<sup>117</sup> It consists of groups of vectographic circles, with one circle appearing at a different depth. The presence of monocular cues in the Titmus Circles was first demonstrated<sup>111</sup> when

10 of 11 (91%) patients with binocular anomalies who could not identify the circle in depth could pass the first level (800") well above chance when asked to identify the circle with a lateral offset. Several studies investigated the presence of monocular cues and all demonstrated a higher-than-chance probability of passing the test monocularly up to the third level (200")<sup>109,112,114,118–121</sup> (at least as per our statistics – see *Tables S1 and S2*; all studies taken together: binomial test with  $n = 472$ ;  $p = 0$ ). The first level (800") could be reached for typical<sup>109,112,113,118,119</sup> or strabismic adults,<sup>112</sup> and typical children<sup>121,122</sup> or children with monofixation syndrome.<sup>114,120</sup> A higher-than-chance proportion of typical children could even pass all levels (800"–40") with one eye patched.<sup>121</sup> In a recent experiment,<sup>83</sup> we demonstrated that binocular non-stereoscopic cues in the Circles provide adults with additional information to pass the test, so that all levels can be passed with a higher-than-chance probability.

We pooled together the data of all previously mentioned studies using the Titmus (*Table S2, Figure 1, n = 647*). Children ( $n = 333$ ) were consistently worse than adults ( $n = 314$ ) suggesting they may be less likely or able to exploit the non-stereoscopic cues in the circles and fly. Observers with pathological vision ( $n = 85$ ) were better than typical observers ( $n = 387$ ) at making use of the



**Figure 1.** Titmus results when the test is administered monocularly, as a function of participants' age (red circles: adults; blue triangles: children) or pathological status (black dots on green background: typical, which means an absence of pathology; crosses on a yellow background: pathological). Ordinate represents the proportion of observers reaching at least the disparity level indicated in abscissa, when pooling together the studies that we reviewed. The dashed black line indicates chance level. Note that probabilities are estimated at the standard Titmus test disparity levels (Circles between 40" and 800", Fly at 3000"). Greater than chance probability indicates the presence of non-stereoscopic cues in the test. Confidence intervals (95%-range) are based on the binomial distribution.

confounding monocular cues present in the largest disparity circle (800") but worse at smaller disparities (40"–140").

### Random Dot E test

After monocular cues were found in the Titmus, random-dot stereograms<sup>123</sup> were introduced in clinical stereo testing. Random-dot stereograms require participants to use global rather than local stereopsis. This strategy was first applied with the Random Dot E<sup>124</sup> (RDE; <https://www.stereo-optical.com/products/stereotests-color-tests/random-dot-e/>), in which the letter E is concealed in one of two plates. However, 94% of observers could pass the RDE test monocularly with one eye or the other.<sup>125</sup> This value is significantly above the chance level of 4.5% for the specific testing procedure they used, clearly demonstrating the presence of monocular cues (binomial test with  $n = 17$ ,  $p = 0$ ).

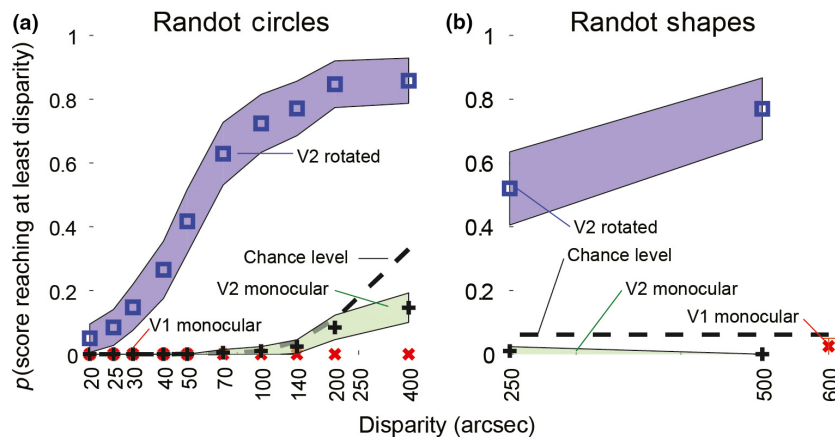
### TNO test

The TNO test<sup>126</sup> (<https://www.good-lite.com/Details.cfm?ProdID=1126>) consists of six random-dot stereogram plates, three for screening with a large disparity and three with fine disparities. The plates depict hidden shapes. The patient's task is to localise and match the shapes. The TNO could not be passed with one eye only (all studies taken together, the proportions were lower than chance level),<sup>109,122,127</sup> but several concerns arise. First, only a red-green anaglyph version of the test is available, which is susceptible to binocular leakage<sup>128</sup> and may trigger suppression.<sup>105,128</sup> Second, as noted earlier, Charman and Jennings<sup>90</sup> identified binocular non-stereoscopic cues by noting that some observers could pass the test when rotated 90°.

### Randot test

The Randot test<sup>129</sup> (<https://www.stereo-optical.com/products/stereotests-color-tests/randot/>) is a clinical stereotest that replaced the Fly portion of the Titmus test with a series of shapes hidden in random-dot stereograms and the Circles portion with a random-dot stereogram version. The two modifications had the objective of removing monocular cues. Indeed, monocular cues were not found for adults<sup>127</sup> nor for children with typical-vision<sup>121,122</sup> or monofixation syndrome.<sup>114</sup> This test has since been discontinued but a later version (version 2) replaced the random-dot stereogram version of the Circles with a contour-based version with dots in the background only.<sup>114</sup> Monocular cues in the Circles were absent or reported less often than chance in three studies with adults.<sup>83,112,113</sup> Two studies<sup>114,120</sup> suggested the presence of monocular cues when testing children with





**Figure 2.** Randot results when its version 1 is administered monocularly (red Xs) and its version 2 is administered either monocularly (black pluses on green background) or rotated (blue squares) for the Circles (a) or the Shapes (b). Ordinate represents the proportion of observers reaching at least the disparity level indicated in abscissa, when pooling together the studies that we reviewed. The black dashed line indicates chance level. Monocular viewing results are at chance, indicating an absence of monocular cues. The rotated version 2 is well above chance indicating the presence of binocular non-stereoscopic cues. Confidence intervals (95%-range) are based on the binomial distribution.

monofixation syndrome monocularly. However, we calculated that the reported proportions are not different from that expected by chance.

Very few adults with typical or strabismic vision<sup>83,112,113</sup> or children with monofixation syndrome<sup>114</sup> could pass the Randot Shapes monocularly, confirming the absence of monocular cues. However, both the Shapes and Circles portions of the test contain binocular non-stereoscopic cues that can be used to pass the test without having stereoscopic vision<sup>83</sup> for observers with typical vision or amblyopia.

We pooled together the data of the previously mentioned studies using the Randot (Table S2, Figure 2,  $n = 536$ ). The two versions were not different from chance when viewed monocularly (proportions were lower than chance). This is to be contrasted with the higher proportion of success as compared to chance level when version 2 is viewed binocularly but rotated ( $n = 93$ ). This is due to binocular non-stereoscopic cues that are available when viewing the test binocularly.<sup>83</sup>

### Butterfly test

The Butterfly stereotest (<https://www.stereoptical.com/products/stereotests-color-tests/butterfly/>) is comprised of the Titmus Circles and a random-dot stereogram depicting the shape of a concealed butterfly. Moll *et al.*<sup>130</sup> recommended the test as an adjunct for screening strabismic patients. The presence of monocular cues in the Butterfly portion has never been investigated, but a recent study<sup>83</sup> demonstrated the presence of non-stereoscopic cues. Seventy-seven percent of the participants reported depth in

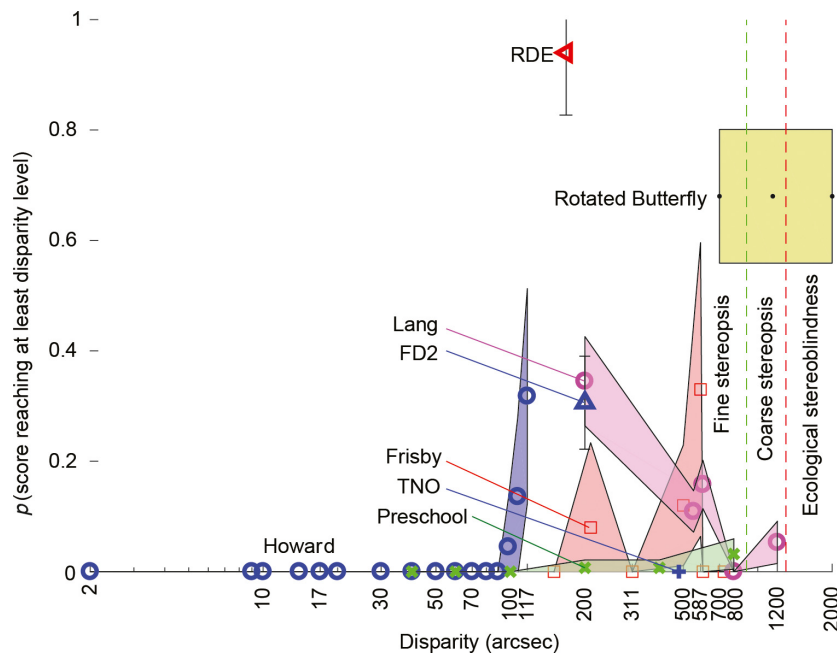
the Butterfly in the rotated viewing condition, showing that rivaldepth occurs (Figure 3).

### Frisby test

Many stereotests require observers to wear either polarised or anaglyphic glasses. Because many children are not compliant, and these glasses were suspected of interocular leakage, Frisby created a natural viewing test of global stereopsis. The Frisby test<sup>131</sup> (<http://frisbystereotest.co.uk/products/frisby-stereotest-near-assesment/>) generates real depth with features printed on each side of two to three Perspex plates of different widths. On one side, the items are organised in a hidden circular shape to localise. However, since the depth is real, it is susceptible to motion parallax cues.<sup>85</sup> Restraining head movements can control this potential confounding cue. Several studies<sup>101,112,113,132,133</sup> reported a probability of patched-eye success no higher than chance which implies that monocular cues are not used to pass the test (see Cooper & Feldman,<sup>132</sup> binomial test with  $n = 34$ ,  $p = 0.09$ ; other studies show lower-than-chance proportions). However, another study reported a proportion of 40%,<sup>109</sup> with an unknown chance level. Taken together (Table S2, Figure 3,  $n = 314$ ), our estimates suggest that the probability of passing the test on non-stereoscopic cues was not significantly different from chance, making this test an attractive possibility.

### Frisby-Davis tests

The Frisby-Davis tests are distance versions of the Frisby test, Frisby-Davis test – the FD2 – and its new version – the



**Figure 3.** Probability of passing the various clinical tests at each tested disparity under non-stereoscopic viewing conditions: monocularly for each test, except for the Butterfly test that is viewed rotated. We pool together all the studies that we reviewed. The tests are the Howard test (blue circles), the Preschool test (green Xs), the TNO test (blue plus), the Frisby test (red squares), the FD2 (blue up-pointing triangle), the Lang test (pink circles), the RDE (red left-pointing triangle) and the Butterfly test viewed rotated (black dots on yellow background). Chance levels are different for each test and therefore not reported on the figure. The vertical green dashed line separates fine and coarse stereopsis, and the vertical red dashed line separates the area of ecological stereoblindness (see below for a definition). Only the Butterfly stereotest shows disparities in that range. Confidence intervals (95%-range) are based on the binomial distribution.

NFD2 (<http://frisbystereotest.co.uk/products/nfd2/>). They are local-stereopsis tests for use at 3–6 m.<sup>134–136</sup> The tests consist of a box containing four rod-mounted shapes, whose individual depth can be changed manually. When tested monocularly, no observers could pass the largest disparity of the FD2.<sup>137</sup> However, in another study,<sup>136</sup> 37% of patients were able to pass the largest disparity with the FD2 at 3 m and 24% at 6 m, both higher-than-chance proportions (binomial tests for  $n = 95$ ,  $p = 10^{-15}$  and  $p = 10^{-6}$ ). Taken together (*Table S2, Figure 3*,  $n = 115$ ), the proportions of success were significantly different from chance. Furthermore, 67% of observers who were stereoblind with the near-Frisby and Randot tests could pass the first level of the FD2 binocularly, indicating non-stereoscopic cues, assuming these patients were genuinely stereoblind. None of these observers obtained a higher score binocularly, which led to the recommendation to attribute stereopsis only if scores were better binocularly than monocularly. Some authors<sup>134</sup> did not find a high probability of success when using that criterion for the NFD2 (3%) or the near-Frisby (1%). The use of such a comparative criterion creates a tangible risk of classifying as stereoblind participants that have both a residual stereovision and a good ability to use potential monocular cues in the test.

### Lang and Lang II tests

The Lang I and Lang II stereotests<sup>138,139</sup> (<https://www.lang-stereotest.com/collections/products>) are random-dot stereograms with concealed shapes to be identified. The Lang I was found to be monocular-cue free for children, with no monocular identification<sup>133</sup> under a procedure for which the chance level was 1.5%. However, a recent study found that many typical-vision adults and strabismic adults could pass several disparity levels of the Lang I and II with one eye covered.<sup>112</sup> While it is likely that these measured proportions of success exceed chance level, the chance level is difficult to calculate when using manufacturer's instructions. Indeed, estimating chance level on this test would require asking hundreds of children for the first three shapes that come to their mind, so as to calculate the frequencies of the Lang's shapes in the responses. The test also requires a shape identification in dense textures, which opens the door to binocular non-stereoscopic cues. Indeed, when strabismic participants could not identify the shape, 61% of them could still localise the area with the larger disparity information for the Lang I, and 51% for the Lang II, suggesting that binocular non-stereoscopic cues are present.<sup>112</sup> Nevertheless, Ancona *et al.*<sup>140</sup> found that of the

four tests (Lang I, Lang II, Titmus, and TNO) they compared to screen for strabismus, the Lang I test had the highest sensitivity, specificity, positive predictive value, and negative predictive value.

### Randot Preschool test

The Randot Preschool Stereotest<sup>141</sup> (<https://www.stereoptical.com/products/stereotests-color-tests/randot-preschool/>) is a test designed for children, using shapes hidden in random-dot stereograms. Manufacturer's instructions (i.e. starting at the 200" disparity) should be avoided because they lead to an inflation of the chance level (47%, see *Table S1*), and instead, the test should be started at the largest disparity, as was done in the following studies. The original version had a defect, a shape with monocular cues at level 800"<sup>113</sup> that was eventually corrected by the manufacturer. For that disparity, the test (excluding the defective shape) resulted in lower-than-chance performance.<sup>113,142</sup> In another study,<sup>112</sup> only 1% of typical-vision observers and 3% of strabismic observers could pass the test with one eye only. Taken together (*Table S2, Figure 3, n = 158*), the proportions were always lower than chance, confirming little if any contributions of monocular cues, making this test an attractive possibility. However, the test task is a shape identification in dense textures, which increases the likelihood that observers use binocular non-stereoscopic cues to pass the test.

### Asteroid

A recent innovative stereotest for children, the Asteroid Test,<sup>143</sup> (<https://research.ncl.ac.uk/asteroid/>) was designed for use without glasses using parallax barrier technology on a 3D tablet. The test uses dynamic random-dot stereograms so that monocular cues cannot be used for disparities of <1000".<sup>144</sup> While some staircase methods may not be a good choice for testing stereopsis, because of the non-monotonic aspect of the psychometric function to disparity, the Asteroid test employs an efficient Bayesian staircase method, which massively reduces the chance level. The Asteroid test is not yet commercially available, but may be a promising resource for future work.

### Laboratory stereotests

We have seen that most clinical stereotests show severe flaws that limit their use for testing stereoblindness in a clinical setting. This is the reason why laboratories often use computer-based psychophysical stereotests. These tests include edge and random-dot stereograms. Edge stereograms (with lines or dots) have several drawbacks. To avoid effective monocular cues, it is possible either (1) to

add large (15') random lateral offsets between the items to compare<sup>33</sup> or, when using pedestals, (2) to randomise convergent and divergent disparities pedestal.<sup>145</sup> When using the first solution, edge stereograms can yield worse acuities than random-dot stereograms because stereo thresholds increase with increasing separation between the items to compare.<sup>26</sup> The second solution leads to the same result because stereo thresholds also increase with increasing depth pedestals.<sup>26</sup> An alternative solution when using gratings is to code the disparity in the carrier with random phases, while the envelope is fixed.<sup>146</sup> However, this procedure introduces a conflict between coarse and fine disparity systems that renders it impractical for testing very large disparities. Finally, when comparing different geometries with vertical offsets between stereograms made of two dots, increased stereo thresholds and variability can also occur.<sup>147</sup>

Monocular cues can be removed using static random-dot stereograms with short presentations ( $\leq 200$  ms) preventing comparisons of offsets between patterns of dots between eyes. However, short presentations can also decrease stereoacuity.<sup>25,45</sup> Alternatively, dynamic random-dot stereograms (i.e., a series of random-dot stereograms depicting the same target/background configuration presented in rapid succession<sup>148</sup>), efficiently prevent monocular cues while allowing for the long presentations needed to maximise stereoacuity. As with edge stereograms, using pedestals and large separations between different targets to compare can decrease stereoacuity: therefore, the recommended task is to compare the target shape with its background. However, random-dot stereograms can introduce binocular non-stereoscopic cues<sup>83</sup> revealing the location and shape of the target on the background. This is because dense textures that are binocularly misaligned can yield rivalry, lustre, rivaldepth or confusion/diplopia. To prevent the observer from using that information, the task should involve localising the depth order between target and background (which one is closer?). In that case, having a clearly visible target shape (e.g. with dots of a different colour) should improve performance and is therefore recommended.

### Recommendation for a quick test for stereoblindness using a clinical stereotest

Based on this review, if only a few minutes are available for testing stereopsis, the Random Dot Butterfly stereotest appears as the best choice if administered as follows. This *modified Butterfly test* makes use exclusively of the Butterfly shape. Observers view the test from a distance of 62 cm, which creates three sets of disparities: the tip of the upper wings are at 1300" (coarse disparities), the bottom wings at 745", and the tip of the abdomen at 450" (fine disparities).

Eleven trials are administered using one of two pairs of polarised glasses, with one pair modified by inverting the left and right eye lenses. After each trial, the observer is asked to change glasses, in a pseudo random fashion so as to test equally with the two glasses. The experimenter/clinician mixes the pairs and ensures the pairs are indistinguishable to the observer, so that the observer does not know which pair is being tested. The task is to report whether the butterfly appears as a hole or as popping out. This modified task instruction is to render binocular non-stereoscopic cues useless. According to our calculations, requiring that observers score 10 out of 11 correct responses in order to pass the test corresponds to  $\approx 91\%$  correct responses with a 0.6%-chance level. To avoid misclassifying inattentive stereo-normal or stereo-weak participants, we recommend running two additional rounds in case of failure in the first round. Three failed rounds would classify the participant as stereoblind. This increases the chance level to 1.8% (expected sensitivity 98.2%) by keeping a high specificity. Such a test may be made of dots too small for some observers. However, it is the only commercially available clinical test that can be easily modified to avoid binocular non-stereoscopic cues while including coarse disparities in the range of ecological stereoblindness. It will be for future studies to validate such a test.

### Prevalence of stereoblindness

What is the prevalence of stereoblindness in the general population? In this section, we will use binomial 95% Confidence Intervals (CI) expressed as percentages. There is a long debate around the question of the prevalence of stereoblindness, with estimations ranging from 1% (CI [0, 6])<sup>14</sup> to 64% (CI [35, 92]).<sup>39</sup> As reviewed in the previous sections, numerous factors are likely at the source of these large variations from the retinal location tested to the contrast of the target used, to cite a few. *Table S3* tabulates the factors that may have affected estimates for each stereoblindness study included in this review.

An unbiased approach to estimate stereoblindness rate is to take the median of all studies (7%; CI [0, 20.5]), which removes the effect of extrema. It is a problematic approach because the result strongly depends on the ratio between studies underestimating and overestimating the true value, which is unknown.

One complexity is that almost all studies used a different criterion for stereoblindness (see *Table S3*), along with additional criteria for stereo-anomaly. For example, Richards<sup>15</sup> reported a stereo-anomaly proportion of 30% (blind to divergent, convergent or zero disparities) but a stereoblindness proportion of only 3% (CI [0, 6]; blind to all disparities). A large-scale study<sup>17</sup> tested over one thousand normal healthy adults. They reported three estimates

of stereoblindness prevalence: 5.3% (CI [4, 6.6]) could not pass a custom circle stereotest, 8.9% (CI [7.2, 10.6]) could not pass the TNO, and 2.2% could not pass any of the tests. While each test measures a different aspect of stereovision (global vs local), each test also allows a different way to hide a potential stereoblindness: the circle test is likely to have monocular cues and the TNO is very likely to have binocular non-stereoscopic cues. It is not clear which of the two reasons can explain the lower last estimate rate.

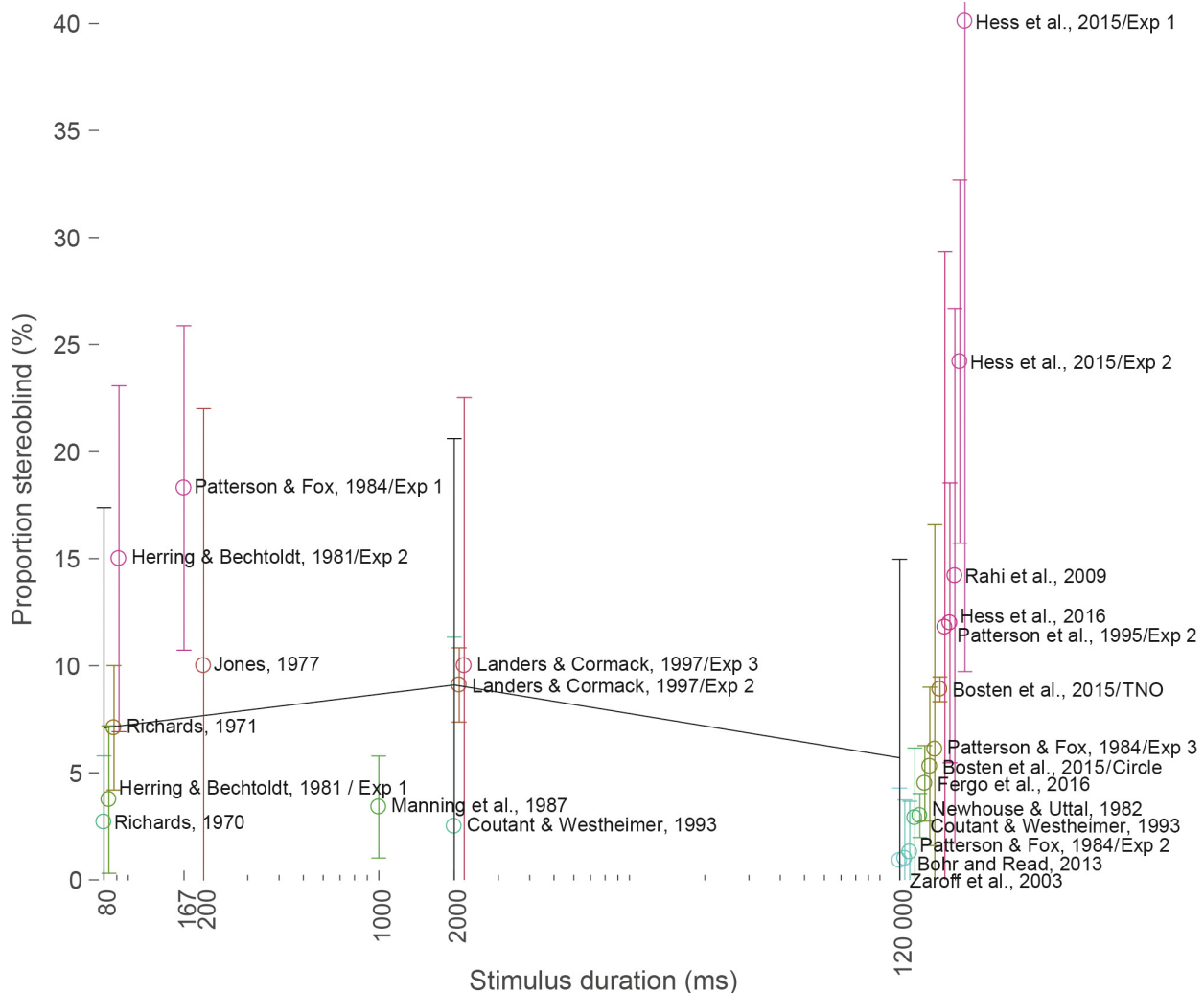
Alternately, the highest estimate for stereoblindness proportion found in the literature – 64% (CI [35, 92])<sup>39</sup> – used a criterion overestimating the real prevalence. Indeed, their criterion for stereoblindness (a gain  $< 0.2$  between depth estimates in binocular and monocular conditions when averaging across all disparities) encompasses all stereo-anomalies, whether related to convergent or divergent disparities. Another influential study also preferred to report the prevalence of stereo-anomaly rather than stereoblindness.<sup>149</sup> In this review, we focus on estimating the prevalence for stereoblindness only. For this reason, we excluded these two last studies in all analyses.

We caution that many studies fail to report critical aspects of their methods. This includes exact stereoblindness criteria, sampling conditions and exclusion criteria. When excluding studies that were ambiguous in their report of their stereoblindness criterion,<sup>150–154</sup> the median stereoblindness rate obtained was 7% (CI [0, 20.5]). Reporting sampling conditions and exclusion criteria is also a particularly important issue, given 91% do not report the way participants were selected.<sup>102</sup> This issue can strongly affect the estimation of stereoblindness prevalence. For example, two studies provide data compatible with very low prevalence. Patterson and Fox<sup>14</sup> reported an incidence of 4% for stereoblindness in adults in their Experiment 2. Problematically, the authors mentioned that participants were recruited from the 30 stereo-anomalous participants detected via their Experiment 1. Only 24 participated, yielding one case of stereoblindness. Assuming the same rate for the six stereo-impaired who did not participate, and taking into account the parent sample ( $n = 98$ ), it can be corrected to a likely 1%-rate (CI [0, 6]). This is probably an underestimation of the true rate because the stereoblindness proportion is unknown in the other 82% of the population who did participate in Experiment 1, when tested under the conditions of Experiment 2. Another study<sup>44</sup> first excluded an unknown fraction of participants with visual defects or a history of it, and Snellen acuity worse than 6/9 (20/30). Among the included participants, 1% of 15-to-59-year-old participants could not report depth in a demo corrugation stimulus and were excluded from further analysis. The rest of the study determines criteria for various levels of stereo-anomaly but only the excluded 1% (CI [0, 2.7]) were considered by the

authors to be stereoblind. The exclusion criteria lead to a strong underestimation of the stereoblindness prevalence in the general population because observers with visual defects are likely to also be stereoblind. When omitting studies that excluded observers based on their visual performances,<sup>14,37,44</sup> the median stereoblindness rate was still 7% (CI [0, 20]).

A different approach to estimating the stereoblindness prevalence is to uncover the study that is theoretically the closest to the true value. For example, two teams<sup>14,149</sup> made the theoretical argument that the large variations in estimated stereoblindness came from differences in the

stimulus duration. According to them, longer stimulus durations are associated with decreased rates of stereo-anomaly. Importantly, this relation was demonstrated for rates of stereo-anomaly but not for rates of stereoblindness. We investigated the relationship between stereoblindness rates and stimulus duration, and did not find the putative relationship (Figure 4; Kruskal-Wallis test  $H = 0.29$ ;  $p = 0.87$ ). Durations <200 ms yielded a median stereoblindness rate of 7% (CI [0, 21]), which increased to 9% (CI [0, 21]) in the range 200–3000 ms. Yet, when presentation were >3000 ms or unlimited, the median rate decreased to 6% (CI [0, 20]). It is possible that the



**Figure 4.** Relation between stereoblindness rates and stimulus duration. Each study is plotted with a circle and a different shade of colour. The continuous black line indicates the medians when pooling together all the reviewed studies in the ranges <200, 200–3000 and >3000 ms (x values are also median of the included studies). Unlimited stimulus duration are given the 120 000 ms value. When more than one study has the same stimulus duration (at 80, 2000, and 120 000 ms), a jitter was added to the x value to prevent super-imposing. Errors bars are 95% confidence intervals based on the binomial distribution. The  $n$  used to calculate the confidence interval of the medians for the different ranges (in black) is the number of studies rather than the number of observers.<sup>156,157</sup>

slightly smaller rate of stereoblindness for very long presentations when compared to very short ones comes from the use of delta-vergence to pass the test with longer stimulus durations.

When pooling together the data from the studies that we reviewed, the range of tested disparities affected the prevalence of stereoblindness (Figure 5). Rates increase from 5.3% (CI [4, 6.6]) for tested ranges that include the smallest disparity (1") to 10% (CI [0, 21]) for tested ranges that include disparities larger than 1° (3600").

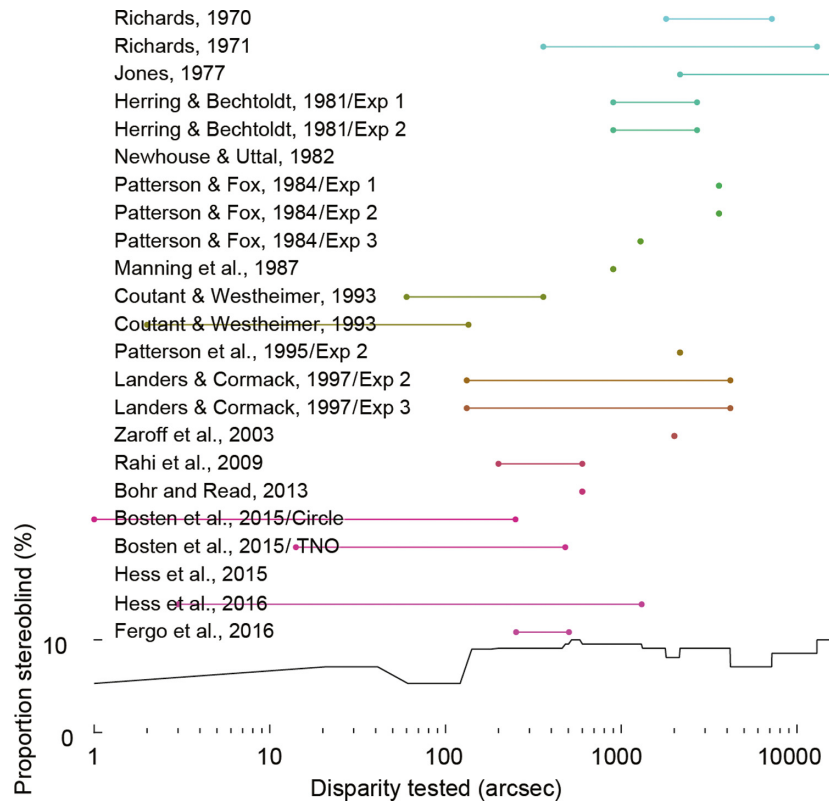
We also considered whether the use of local stereograms may affect the result. There are two reasons they could do so: (1) because local stereograms are thought to be easier than global stereograms, and (2) because local stereograms have monocular cues that can be used to deceive the test. Although we found a trend for rates to be twice lower when tested with local stereograms than with global stereograms (medians: 6% vs 10%), the distributions were not statistically different (Mann-Whitney rank-sum = 131;  $p = 0.29$ ). Monocular cues also had little influence on the results. We classified each study according to the likelihood to show useful monocular cues to pass the test (Table S3): likely or unlikely. The different classes were not significantly

associated with different stereoblindness rates (Kruskal-Wallis  $H = 1.3, p = 0.52$ ).

Finally, we also classified the studies according to the likelihood of presence of binocular non-stereoscopic cues (likely or unlikely). The different classes were not significantly associated with different stereoblindness rates (Kruskal-Wallis  $H = 0.004, p = 0.95$ ).

Another approach is to select studies that match multiple important criteria. The criteria we have chosen are studies that reported their stereoblindness criterion unambiguously and tested at least a fine disparity (<900") and a coarse disparity (>900"). Only one study qualified<sup>43</sup> and yielded a stereoblindness rate of 7% (CI [0, 21]).

Four different approaches converged toward a prevalence of stereoblindness of 7% (median approach: 7%; unambiguous-stereoblindness-criteria approach: 7%; visual-defect-included approach: 7%; multiple-criteria approach: 7%). It should be noted that some of the studies include adults of age >60 years and that we have excluded studies focusing exclusively on adults >60 years. Older adults usually have a higher incidence of stereo-anomaly.<sup>44,134,154</sup> Future studies could map out how the prevalence changes with maturation and ageing. Finally, a refined estimate



**Figure 5.** Relation between stereoblindness rate and tested disparity. The range of tested disparities for each study is plotted for each study: the minimal and maximal tested disparities are indicated with two dots that are connected. If only one disparity was tested, only a dot is displayed. The bottom black line depicts the median of the stereoblind rates found by all studies that have tested disparities in the range  $[x, x + 10]$ .

would use the same procedure with observers with visual pathologies. While some of these studies specifically exclude individuals with amblyopia, it is not clear whether others included individuals with amblyopia or with strabismus without amblyopia.

### An ecological definition for stereoblindness

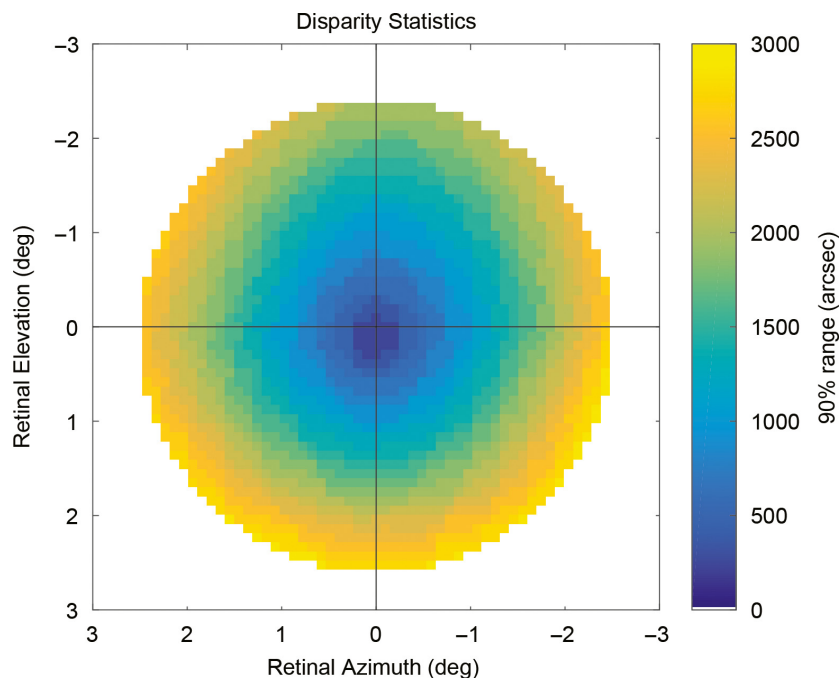
There is presently no simple definition for stereoblindness. Most basic and clinical studies have used one or more commercial stereo tests, and assigned either 'nil' stereo acuity, a stereo sensitivity of zero, or an arbitrary large disparity (often twice the largest disparity tested) to those who fail to pass the test. However, as noted in the earlier sections, an individual may fail a particular test for any one of many reasons (e.g., because the disparities are not large enough; the stimulus may not be large enough; and/or the contrast may be too low). In our earlier review,<sup>12</sup> we found that  $\approx 27\%$  of patients who had 'nil' stereopsis prior to training had some measurable stereoacuity following training. However, most of those studies used clinical tests that may not have revealed the patient's 'true' baseline stereoacuity, or their capacity for stereovision. Indeed, it is possible that stereoblindness, like blindness, is a concept and not necessarily a physiological state. Patients with visual acuity worse than 6/60 (20/200) are considered legally blind; however,

most 'blind' patients have at least some light perception – they can often tell light from dark. Thus, it is possible that many patients labelled 'stereoblind' may actually have at least some rudimentary stereovision under the right conditions.

How should we define stereoblindness? We propose an ecological definition based on the inability to detect disparities within the range of disparities that are found in natural scenes.<sup>155</sup> Specifically, across fixation distances and scenes (both indoor and outdoor), 90% of the disparities within the fovea ( $\pm 2.5^\circ$  radius) are within  $\pm 1314''$  (Figure 6). Therefore, we propose that ecological stereoblindness be defined as a stereoacuity of  $>1300''$ , and that total stereoblindness be defined as no stereo perception at all in any condition. The proposed criterion of stereoacuity  $>1300''$  is larger than the  $1040''$  average maximum disparity found in clinical tests (Table S1). Given this definition, a rapid way to screen for stereoblindness would be to present our modified Butterfly test at 62 cm, which would give the wing tips a disparity of  $1300''$ .

### Summary and conclusions

We reviewed the optimal conditions for measuring stereoacuity with the goal of detecting stereoblindness. We showed how variations in measuring stereoblindness



**Figure 6.** The fifth to ninety-fifth percentile of disparities present in natural scenes (across viewing distances and tasks) in the central  $5^\circ$ . Azimuth and Elevation are expressed in degrees, in a Helmholtz reference frame. Positive Azimuth is left and positive Elevation is down. Crossed disparity is positive, uncrossed disparity is negative. Figure courtesy of Martin Banks and Agostino Gibaldi (based on data from Sprague, et al.<sup>155</sup>).

influence the prevalence of the defect and proposed corrections to extract the most accurate estimate. Following different approaches, we estimate the prevalence of stereoblindness to be 7%. We note that these estimates are heavily weighted to adults of age <60 years old, with older adults possibly having a higher prevalence. Finally, we make recommendations for a new ecological definition for stereoblindness and for efficient clinical methods for testing stereoacuity and determining stereoblindness.

These considerations should serve as helpful guide for intervention studies looking into reducing stereo-impairment or getting stereoblind participants to recover some stereopsis. Understanding the extent to which stereoscopic vision is plastic and thus trainable relies indeed critically on our ability to properly estimate stereo-impairments and stereoblindness.

## Acknowledgements

This work was supported by grant number RO1EY020976 from the National Eye Institute. We thank Martin Banks and Agostino Gibaldi for their help in providing ecological data and material, and Jenny Read and an anonymous referee for their insightful comments and suggestions on an earlier version of this paper.

## Disclosure

The authors report no conflicts of interest and have no proprietary interest in any of the materials mentioned in this article.

## References

- McKee SP & Taylor DG. The precision of binocular and monocular depth judgments in natural settings. *J Vis* 2010; 10: 5.
- Howard CHJ. A test for the judgment of distance. *Am J Ophthalmol* 1919; 2: 656–675.
- Caziot B & Backus BT. Stereoscopic offset makes objects easier to recognize. *PLoS One* 2015; 10: e0129101.
- Bradshaw M, Elliott K, Hibbard P, Simpson PJ, Watt S & Davies I. Binocular cues and the control of prehension. *Spat Vis* 2004; 17: 95–110.
- Vergheze P, Tyson TL, Ghahghaei S & Fletcher DC. Depth perception and grasp in central field loss. *Invest Ophthalmol Vis Sci* 2016; 57: 1476–1487.
- Knill DC. Reaching for visual cues to depth: the brain combines depth cues differently for motor control and perception. *J Vis* 2005; 5: 103–115.
- Jones RK & Lee DN. Why two eyes are better than one: the two views of binocular vision. *J Exp Psychol Hum Percept Perform* 1981; 7: 30–40.
- Sheedy JE, Bailey IL, Buri M & Bass E. Binocular vs. monocular task performance. *Am J Optom Physiol Opt* 1986; 63: 839–846.
- Barry SR. *Fixing my Gaze: A Scientist's Journey Into Seeing in Three Dimensions*. Basic Books: New York, NY, 2009.
- Holmes G & Horrax G. Disturbances of spatial orientation and visual attention, with loss of stereoscopic vision. *Arch Neurol Psychiatry* 1919; 1: 384–407.
- Bridge H. Effects of cortical damage on binocular depth perception. *Philos Trans R Soc Lond B Biol Sci* 2016; 371: 20150254.
- Levi DM, Knill DC & Bavelier D. Stereopsis and amblyopia: a mini-review. *Vision Res* 2015; 114: 17–30.
- Worth C & Chavasse FB. Isolated paralysis of the medial rectus. In: *Squint*, Lyle TK (ed.). Blakiston: Philadelphia, PA, 1950; p. 106.
- Patterson R & Fox R. The effect of testing method on stereoanomaly. *Vision Res* 1984; 24: 403–408.
- Richards W. Stereopsis and stereoblindness. *Exp Brain Res* 1970; 10: 380–388.
- Rahi JS, Cumberland PM & Peckham CS. Visual impairment and vision-related quality of life in working-age adults: findings in the 1958 British birth cohort. *Ophthalmology* 2009; 116: 270–274.
- Bosten JM, Goodbourn PT, Lawrance-Owen AJ, Bargary G, Hogg R & Mollon JD. A population study of binocular function. *Vision Res* 2015; 110: 34–50.
- Haegerstrom-Portnoy G, Schneck ME & Brabyn JA. Seeing into old age: vision function beyond acuity. *Optom Vis Sci* 1999; 76: 141–158.
- Lakshmanan Y & George RJ. Stereoacuity in mild, moderate and severe glaucoma. *Ophthalmic Physiol Opt* 2013; 33: 172–178.
- Laidlaw DAH, Hammad RA, Hopper CD *et al*. Randomised trial of effectiveness of second eye cataract surgery. *Lancet* 1998; 352: 925–929.
- Elliott DB, Patla AE, Furniss M & Adkin A. Improvements in clinical and functional vision and quality of life after second eye cataract surgery. *Optom Vis Sci* 2000; 77: 13–24.
- Simmerman JS. Absolute threshold of stereopsis using the Frisby stereotest. *J Am Optom Assoc* 1984; 55: 50–53.
- Westheimer G. Clinical evaluation of stereopsis. *Vision Res* 2013; 90: 38–42.
- Legge GE & Yuanchao G. Stereopsis and contrast. *Vision Res* 1989; 29: 989–1004.
- Westheimer G & Pettet MW. Contrast and duration of exposure differentially affect vernier and stereoscopic acuity. *Proc R Soc B Biol Sci* 1990; 241: 42–46.
- Blakemore C. The range and scope of binocular depth discrimination in man. *J Physiol* 1970; 211: 599–622.
- Westheimer G & McKee S. Stereogram design for testing local stereopsis. *Invest Ophthalmol Vis Sci* 1980; 19: 802–809.
- Woo GC & Sillanpaa V. Absolute stereoscopic thresholds as measured by crossed and uncrossed disparities. *Am J Optom Physiol Opt* 1979; 56: 350–355.



29. McKee SP, Levi DM & Bowne SF. The imprecision of stereopsis. *Vision Res* 1990; 30: 1763–1779.
30. Fahle M, Edelman S & Poggio T. Fast perceptual learning in hyperacuity. *Vision Res* 1995; 35: 3003–3013.
31. Fendick M & Westheimer G. Effects of practice and the separation of test targets on foveal and peripheral stereoacuity. *Vision Res* 1983; 23: 145–150.
32. McKee SP, Welch L, Taylor DG & Bowne SF. Finding the common bond: Stereoacuity and the other hyperacuities. *Vision Res* 1990; 30: 879–891.
33. Ding J & Levi DM. Recovery of stereopsis through perceptual learning in human adults with abnormal binocular vision. *Proc Natl Acad Sci USA* 2011; 108: E733–E741.
34. Sireteanu R, Fronius M & Singer W. Binocular interaction in the peripheral visual field of humans with strabismic and anisometropic amblyopia. *Vision Res* 1981; 21: 1065–1074.
35. Wilmer JB. How to use individual differences to isolate functional organization, biology, and utility of visual functions; with illustrative proposals for stereopsis. *Spat Vis* 2008; 21: 561–579.
36. Mustillo P. Binocular mechanisms mediating crossed and uncrossed stereopsis. *Psychol Bull* 1985; 97: 187.
37. Landers DD & Cormack LK. Asymmetries and errors in perception of depth from disparity suggest a multicomponent model of disparity processing. *Percept Psychophys* 1997; 59: 219–231.
38. Richards W & Kaye MG. Local versus global stereopsis: two mechanisms? *Vision Res* 1974; 14: 1345–1347.
39. Van Ee R & Richards W. A planar and a volumetric test for stereoaomaly. *Perception* 2002; 31: 51–64.
40. Dorman R & van Ee R. 50 years of stereoblindness: reconciliation of a continuum of disparity detectors with blindness for disparity in near or far depth. *Iperception* 2017; 8: 2041669517738542.
41. Jones R. Anomalies of disparity detection in the human visual system. *J Physiol* 1977; 264: 621–640.
42. Patterson R, Cayko R, Short GL et al. Temporal integration differences between crossed and uncrossed stereoscopic mechanisms. *Percept Psychophys* 1995; 57: 891–897.
43. Richards W. Anomalous stereoscopic depth perception. *J Opt Soc Am* 1971; 61: 410–414.
44. Zaroff CM, Knutelska M & Frumkes TE. Variation in stereoacuity: normative description, fixation disparity, and the roles of aging and gender. *Invest Ophthalmol Vis Sci* 2003; 44: 891–900.
45. Patterson R & Martin WL. Human stereopsis. *Hum Factors* 1992; 34: 669–692.
46. Harris JM & Parker AJ. Independent neural mechanisms for bright and dark information in binocular stereopsis. *Nature* 1995; 374: 808–811.
47. Read JCA, Vaz XA & Serrano-Pedraza I. Independent mechanisms for bright and dark image features in a stereo correspondence task. *J Vis* 2011; 11: 4.
48. Read JCA & Cumming BG. A stimulus artefact undermines the evidence for independent ON and OFF channels in stereopsis. *BioRxiv* 2018. <https://doi.org/10.1101/295618>
49. Richards W. Reversal in stereo discrimination by contrast reversal. *Am J Optom Arch Am Acad Optom* 1973; 50: 968–977.
50. Zinn WJ & Solomon H. A comparison of static and dynamic stereoacuity. *J Am Optom Assoc* 1985; 56: 712–715.
51. Regan D, Erkelens CJ & Collewijn H. Visual field defects for vergence eye movements and for stereomotion perception. *Invest Ophthalmol Vis Sci* 1986; 27: 806–819.
52. Kitaoji H & Toyama K. Preservation of position and motion stereopsis in strabismic subjects. *Invest Ophthalmol Vis Sci* 1987; 28: 1260–1267.
53. Watanabe Y, Kezuka T, Harasawa K, Usui M, Yaguchi H & Shioiri S. A new method for assessing motion-in-depth perception in strabismic patients. *Br J Ophthalmol* 2008; 92: 47–50.
54. Nefs HT, O'Hare L & Harris JM. Two independent mechanisms for motion-in-depth perception: evidence from individual differences. *Front Psychol* 2010; 1: 155.
55. Hess RF, Mansouri B, Thompson B & Gheorghiu E. Latent stereopsis for motion in depth in strabismic amblyopia. *Invest Ophthalmol Vis Sci* 2009; 50: 5006–5016.
56. Westheimer G & McKee SP. What prior uniocular processing is necessary for stereopsis? *Invest Ophthalmol Vis Sci* 1979; 18: 614–621.
57. Chopin A, Levi DM, Knill D & Bavelier D. The absolute disparity anomaly and the mechanism of relative disparities. *J Vis* 2016; 16: 1–17.
58. Read JCA, Phillipson GP, Serrano-Pedraza I, Milner AD & Parker AJ. Stereoscopic vision in the absence of the lateral occipital cortex. *PLoS One* 2010; 5: e12608.
59. Ogle KN. Disparity limits of stereopsis. *Arch Ophthalmol* 1952; 48: 50–60.
60. Ogle KN. On the limits of stereoscopic vision. *J Exp Psychol* 1952; 44: 253–259.
61. Ogle KN. Precision and validity of stereoscopic depth perception from double images. *J Opt Soc Am* 1953; 43: 906–913.
62. Wilcox LM & Allison RS. Coarse-fine dichotomies in human stereopsis. *Vision Res* 2009; 49: 2653–2665.
63. Giaschi D, Narasimhan S, Solski A, Harrison E & Wilcox LM. On the typical development of stereopsis: fine and coarse processing. *Vision Res* 2013; 89: 65–71.
64. Giaschi D, Lo R, Narasimhan S, Lyons C & Wilcox LM. Sparing of coarse stereopsis in stereodeficient children with a history of amblyopia. *J Vis* 2013; 13: 17.
65. Shimono K. Evidence for the subsystems in stereopsis: fine and coarse stereopsis. *Jpn Psychol Res* 1984; 26: 168–172.
66. Palmer DA. Measurement of the horizontal extent of Panum's area by a method of constant stimuli. *Opt Acta Int J Opt* 1961; 8: 151–159.

67. Levitt H. Transformed up-down methods in psychoacoustics. *J Acoust Soc Am* 1971; 49: 467–477.
68. Wilcox LM & Hess RF. Dmax for stereopsis depends on size, not spatial frequency content. *Vision Res* 1995; 35: 1061–1069.
69. Schor CM, Edwards M & Pope DR. Spatial-frequency and contrast tuning of the transient-stereopsis system. *Vision Res* 1998; 38: 3057–3068.
70. Westheimer G & Tanzman IJ. Qualitative depth localization with diplopic images. *J Opt Soc Am* 1956; 46: 116–117.
71. Pope DR, Edwards M & Schor CS. Extraction of depth from opposite-contrast stimuli: transient system can, sustained system can't. *Vision Res* 1999; 39: 4010–4017.
72. Hess RF & Wilcox LM. Linear and non-linear filtering in stereopsis. *Vision Res* 1994; 34: 2431–2438.
73. Hess RF & Wilcox LM. The transient nature of 2nd-order stereopsis. *Vision Res* 2008; 48: 1327–1334.
74. Julesz B & Spivack GJ. Stereopsis based on vernier acuity cues alone. *Science* 1967; 157: 563–565.
75. Rose D & Price E. Functional separation of global and local stereopsis investigated by cross-adaptation. *Neuropsychologia* 1995; 33: 269–274.
76. Ross JE. Disturbance of stereoscopic vision in patients with unilateral stroke. *Behav Brain Res* 1983; 7: 99–112.
77. Hamsher KD. Stereopsis and unilateral brain disease. *Invest Ophthalmol Vis Sci* 1978; 17: 336–343.
78. Ptito A, Zatorre RJ, Larson WL & Tosoni C. Stereopsis after unilateral anterior temporal lobectomy: dissociation between local and global measures. *Brain* 1991; 114: 1323–1333.
79. Read JCA. Stereo vision and strabismus. *Eye* 2015; 29: 214–224.
80. Enright JT. Exploring the third dimension with eye movements: better than stereopsis. *Vision Res* 1991; 31: 1549–1562.
81. Backus BT & Matza-Brown D. The contribution of vergence change to the measurement of relative disparity. *J Vis* 2003; 3: 737–750.
82. Masson GS, Busetini C & Miles FA. Vergence eye movements in response to binocular disparity without depth perception. *Nature* 1997; 389: 283–286.
83. Chopin A, Chan SW, Guellaï B, Bavelier D & Levi DM. *Deceiving clinical stereotests with a new family of cues: binocular non-stereoscopic cues*. [https://figshare.com/articles/Deceiving\\_clinical\\_stereotests\\_with\\_a\\_new\\_family\\_of\\_cues\\_s\\_binocular\\_non-stereoscopic\\_cues/7152626](https://figshare.com/articles/Deceiving_clinical_stereotests_with_a_new_family_of_cues_s_binocular_non-stereoscopic_cues/7152626) (accessed 23/1/2019).
84. Cutting JE & Vishton PM. Perceiving layout and knowing distances: The integration, relative potency, and contextual use of different information about depth. In: *Handbook of Perception and Cognition*, Epstein W & Rogers S (eds). Academic Press: San Diego, CA, 1995; pp. 69–117.
85. Rogers B & Graham M. Motion parallax as an independent cue for depth perception. *Perception* 1979; 8: 125–134.
86. Held RT, Cooper EA & Banks MS. Blur and disparity are complementary cues to depth. *Curr Biol* 2012; 22: 426–431.
87. Archer SM. Stereotest artifacts and the strabismus patient. *Graefes Arch Clin Exp Ophthalmol* 1988; 226: 313–316.
88. O'Connor AR & Tidbury LP. Stereopsis: are we assessing it in enough depth? *Clin Exp Optom* 2018; 101: 485–494.
89. Cooper J. Clinical stereopsis testing: contour and random dot stereograms. *J Am Optom Assoc* 1979; 50: 41–46.
90. Charman WN & Jennings JAM. Recognition of TNO stereotest figures in the absence of true stereopsis. *Optom Vis Sci* 1995; 72: 535–536.
91. Meenes M. A phenomenological description of retinal rivalry. *Am J Psychol* 1930; 42: 260–269.
92. O'Shea RP & Blake R. Depth without disparity in random-dot stereograms. *Percept Psychophys* 1987; 42: 205–214.
93. Wolfe JM. Influence of spatial frequency, luminance, and duration on binocular rivalry and abnormal fusion of briefly presented dichoptic stimuli. *Perception* 1983; 12: 447–456.
94. Schor CM. Visual stimuli for strabismic suppression. *Perception* 1977; 6: 583–593.
95. Baitch LW & Levi DM. Evidence for nonlinear binocular interactions in human visual cortex. *Vision Res* 1988; 28: 1139–1143.
96. Guclu H, Gurlu VP, Ozal SA & Ozkurt ZG. Prognostic factors for stereopsis in refractive accommodative esotropia. *Pakistan J Med Sci* 2015; 31: 807–811.
97. Andalib D, Nabie R & Poormohammad B. Factors affecting improvement of stereopsis following successful surgical correction of childhood strabismus in adults. *Strabismus* 2015; 23: 80–84.
98. Ohlsson J, Villarreal G, Abrahamsson M, Cavazos H, Sjöström A & Sjöstrand J. Screening merits of the Lang II, Frisby, Randot, Titmus, and TNO stereo tests. *J Am Assoc Pediatr Ophthalmol Strabismus* 2001; 5: 316–322.
99. Farvardin M & Afarid M. Evaluation of stereo tests for screening of amblyopia. *Iran Red Crescent Med J* 2007; 9: 80–85.
100. Ohlsson J, Villarreal G, Sjöström A, Abrahamsson M & Sjöstrand J. Screening for amblyopia and strabismus with the Lang II stereo card. *Acta Ophthalmol Scand* 2002; 80: 163–166.
101. Fricke TR & Siderov J. Stereopsis, stereotests, and their relation to vision screening and clinical practice. *Clin Exp Optom* 1997; 80: 165–172.
102. Heron S & Lages M. Screening and sampling in studies of binocular vision. *Vision Res* 2012; 62: 228–234.
103. Howard IP & Rogers BJ. Tests of stereoscopic vision. In: *Perceiving in Depth, Volume 2: Stereoscopic Vision*, Howard IP & Rogers BJ (eds.). Oxford University Press Inc: New York, NY, 2012; pp. 289–295.
104. Pearlman JT. Stereoscopic vision testing: a review. *Am Orthopt J* 1969; 19: 78–86.
105. Hincliffe HA. Clinical evaluation of stereopsis. *Br Orthopt J* 1978; 35: 46–57.
106. Adams WE, Leske DA, Hatt SR & Holmes JM. Defining real change in measures of stereoacuity. *Ophthalmology* 2009; 116: 281–285.

107. Sloan LL & Altman A. Factors involved in several tests of binocular depth perception. *AMA Arch Ophthalmol* 1954; 52: 524–544.
108. Woodburne LS. The effect of a constant visual angle upon the binocular discrimination of depth differences. *Am J Psychol* 1934; 46: 273–286.
109. Hall C. The relationship between clinical stereotests. *Ophthalmic Physiol Opt* 1982; 2: 135–143.
110. Feinberg R & Reuel S. *Device for testing visual acuity [Internet]*. US3011394 A, 1961.
111. Simons K & Reinecke RD. A reconsideration of amblyopia screening and stereopsis. *Am J Ophthalmol* 1974; 78: 707–713.
112. Hahn E, Comstock D, Durling S et al. Monocular clues in seven stereotests. *Dalhousie Med J* 2010; 37: 5–13.
113. Holmes JM & Leske DA. Monocular clues in tests of stereoacuity. In: *Transactions IX International Orthoptic Congress*. Berufsverband der Orthoptistinnen Deutschlands e. V: Nurnberg, Germany, 1999; pp. 103–106.
114. Schweers MA & Baker JD. Comparison of Titmus® and two Randot™ tests in monofixation. *Am Orthopt J* 1992; 42: 135–141.
115. Arnoldi K & Frenkel A. Modification of the Titmus fly test to improve accuracy. *Am Orthopt J* 2014; 64: 64–70.
116. De La Cruz A, Morale SE, Jost RM, Kelly KR & Birch EE. Modified test protocol improves sensitivity of the stereo fly test. *Am Orthopt J* 2016; 66: 122–125.
117. Fawcett SL. An evaluation of the agreement between contour-based circles and random dot-based near stereoacuity tests. *Am Assoc Pediatr Ophthalmol Strabismus* 2005; 9: 572–578.
118. Cooper J & Warshowsky J. Lateral displacement as a response cue in the Titmus stereo test. *Am J Optom Physiol Opt* 1977; 54: 537–541.
119. Levy NS & Glick EB. Stereoscopic perception and Snellen visual acuity. *Am J Ophthalmol* 1974; 78: 722–724.
120. Clarke WN & Noel LP. Stereoacuity testing in the monofixation syndrome. *J Pediatr Ophthalmol Strabismus* 1990; 27: 161–163.
121. Cooper J, Feldman J & Medlin D. Comparing stereoscopic performance of children using the Titmus, TNO, and Randot stereo tests. *J Am Optom Assoc* 1979; 50: 821–825.
122. Brown MH & Morris JE. Comparative analysis of current stereotests. *Am Orthopt J* 1980; 30: 93–96.
123. Julesz B. Binocular depth perception of computer-generated patterns. *Bell Syst Tech J* 1960; 39: 1125–62.
124. Reinecke RD & Simons K. A new stereoscopic test for amblyopia screening. *Am J Ophthalmol* 1974; 78: 714–721.
125. Fricke TR & Siderov J. Non-stereoscopic cues in the Random-Dot E stereotest: results for adult observers. *Ophthalmic Physiol Opt* 1997; 17: 122–127.
126. Walraven J. Amblyopia Screening with random-dot stereograms. *Am J Ophthalmol* 1975; 80: 893–900.
127. Mazow ML, Prager TC & Cathey G. Assessment of three stereo acuity tests. *Am Orthopt J* 1983; 33: 111–115.
128. Simons K & Elhatton K. Artifacts in fusion and stereopsis testing based on red/green dichoptic image separation. *J Pediatr Ophthalmol Strabismus* 1994; 31: 290–297.
129. Alexander JA. A New clinical test of stereopsis: theoretical evaluation. *Aust J Optom* 1979; 62: 191–193.
130. Moll AM, Rao RC, Rotberg LB, Roarty JD, Bohra LI & Baker JD. The role of the random dot Stereo Butterfly test as an adjunct test for the detection of constant strabismus in vision screening. *J Am Assoc Pediatr Ophthalmol Strabismus* 2009; 13: 354–356.
131. Sasieni LS. The Frisby stereotest. *Optician* 1978; 176: 7–10.
132. Cooper J & Feldman J. Assessing the Frisby stereo test under monocular viewing conditions. *J Am Optom Assoc* 1979; 50: 807–809.
133. Manny RE, Martinez AT & Fern KD. Testing stereopsis in the preschool child: is it clinically useful? *J Pediatr Ophthalmol Strabismus* 1991; 28: 223–231.
134. Bohr I & Read JCA. Stereoacuity with Frisby and revised FD2 stereo tests. *PLoS One* 2013; 8: e82999.
135. Frisby JP & Davis H. Clinical tests of distance stereopsis: state of the art. In: *Int Strabismological Assoc ISA 2002*, de Faber J-T (ed.). Swets & Zeitlinger: Lisse, 2003; pp. 187–190.
136. Holmes JM & Fawcett SL. Testing distance stereoacuity with the Frisby-Davis 2 (FD2) test. *Am J Ophthalmol* 2005; 139: 193–195.
137. Laird PW, Hatt SR, Leske DA & Holmes JM. Distance stereoacuity in prism-induced convergence stress. *J Am Assoc Pediatr Ophthalmol Strabismus* 2008; 12: 370–374.
138. Lang J. A new stereotest. *J Pediatr Ophthalmol Strabismus* 1983; 20: 72–74.
139. Nüssgens Z, Czerwonka B & Roggenkämper P. Examinations on the new Lang test. *Strabismus* 1993; 1: 69–73.
140. Ancona C, Stoppani M, Odazio V, La Spina C, Corradetti G & Bandello F. Stereo tests as a screening tool for strabismus: which is the best choice? *Clin Ophthalmol* 2014; 8: 2221–2227.
141. Birch E, Williams C, Hunter J & Lapa MC. Random dot stereoacuity of preschool children. ALSPAC “Children in Focus” Study Team. *J Pediatr Ophthalmol Strabismus* 1997; 34: 217–222.
142. Leske DA & Holmes JM. Maximum angle of horizontal strabismus consistent with true stereopsis. *J Am Assoc Pediatr Ophthalmol Strabismus* 2004; 8: 28–34.
143. Read JCA. What is stereoscopic vision good for? In: *Proc SPIE 9391, Stereoscopic Displays and Applications XXVI*, Holliman NS, Woods AJ, Favalora GE & Kawai T (eds). International Society for Optics and Photonics: Bellingham, WA, USA, 2015; p. 93910N.
144. Serrano-Pedraza I, Vancleef K & Read JCA. Avoiding monocular artifacts in clinical stereotests presented on column-interleaved digital stereoscopic displays. *J Vis* 2016; 16: 13.

145. Siderov J & Harwerth RS. Precision of stereoscopic depth perception from double images. *Vision Res* 1993; 33: 1553–1560.
146. Farell B, Li S & McKee SP. Disparity increment thresholds for gratings. *J Vis* 2004; 4: 156–168.
147. Harris JM, Chopin A, Zeiner KM & Hibbard PB. Perception of relative depth interval: systematic biases in perceived depth. *Q J Exp Psychol* 2012; 65: 73–91.
148. Breitmeyer B, Julesz B & Kropfl W. Dynamic random-dot stereograms reveal up-down anisotropy and left-right isotropy between cortical hemifields. *Science* 1975; 187: 269–270.
149. Tam WJ & Stelmach LB. Display duration and stereoscopic depth discrimination. *Can J Exp Psychol* 1998; 52: 56–61.
150. Manning ML, Finlay DC, Neill RA & Frost BG. Detection threshold differences to crossed and uncrossed disparities. *Vision Res* 1987; 27: 1683–1686.
151. Coutant BE & Westheimer G. Population distribution of stereoscopic ability. *Ophthalmic Physiol Opt* 1993; 13: 3–7.
152. Hess RF, To L, Zhou J, Wang G & Cooperstock JR. Stereo vision: the haves and have-nots. *Iperception* 2015; 6: 2041669515593028.
153. Hess RF, Ding R, Clavagnier S et al. A robust and reliable test to measure stereopsis in the clinic. *Invest Ophthalmol Vis Sci* 2016; 57: 798.
154. Fergo C, Burcharth J, Pommergaard H-C & Rosenberg J. Age is highly associated with stereo blindness among surgeons: a cross-sectional study. *Surg Endosc* 2016; 30: 4889–4894.
155. Sprague WW, Cooper EA, Tošić I & Banks MS. Stereopsis is adaptive for the natural environment. *Sci Adv* 2015; 1: e1400254.
156. Herring RD & Bechtoldt HP. Categorical perception of stereoscopic stimuli. *Percept Psychophys* 1981; 29: 129–137.
157. Newhouse M & Uttal WR. Distribution of stereoanomalies in the general population. *Bull Psychon Soc* 1982; 20: 48–50.

### Supporting Information

Additional Supporting Information may be found in the online version of this article:

**Table S1.** The different clinical stereopsis tests.

**Table S2.** Results when testing the presence of non- stereoscopic cues.

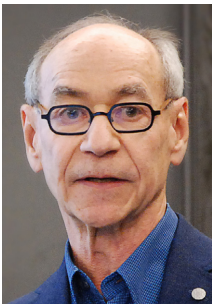
**Table S3.** Factors that may have affected estimates of stereoblindness.



**Adrien Chopin** is a research and teaching fellow in psychology and neurosciences within the University of Geneva in Switzerland. He received his PhD from Paris Descartes University and then conducted research mostly in the University of Geneva and the University of California, Berkeley. His PhD was on probabilistic inferences in the binocular visual system. His current work aims at understanding binocular vision and binocular pathologies (stereoblindness and amblyopia) using psychophysics and fMRI. He conducts remediation research through training that triggers learning (stereo-recovery, video gaming, virtual reality), and he studies brain plasticity with fMRI.



**Daphne Bavelier** studies how the brain adapts to changes in experience, either by nature - for example, deafness - or by training - for example, playing video games. Initially trained in Biology at the 'Ecole Normale Supérieure de Paris', she then received a PhD in Brain and Cognitive Sciences from MIT and trained in human brain plasticity at the Salk Institute. Her work shows that playing fast-paced, action-packed entertainment video games typically thought to be mind-numbing actually benefits several aspects of behavior. Exploiting this counter-intuitive finding, her lab now investigates how new media, such as video games, can be leveraged to foster learning and brain plasticity.



**Dennis Michael Levi** is a Professor at the University of California, Berkeley, with appointments in the School of Optometry, The Graduate Group in Vision Science, and the Helen Wills Neuroscience Institute. His research focuses on how we see form and depth, and how these are impacted by abnormal early visual development, particularly amblyopia and strabismus. Levi's research has been funded by the National Eye Institute (NEI) since 1976. He has received numerous awards including the Glenn Fry Award, the Garland Clay Award and the Prentice Medal from the American Academy of Optometry, and most recently the 2016 Edgar D. Tillyer Award from the Optical Society of America, for his groundbreaking work on treating amblyopia.