Journal of Cognitive Enhancement
https://doi.org/10.1007/s41465-024-00299-w

ORIGINAL RESEARCH

PLFest: A New Platform for Accessible, Reproducible, and Open Perceptual Learning Research

Samyukta Jayakumar[1] · Marcello Maniglia[1] · Zhenhong Guan[2] · C. Shawn Green[2] · Aaron R. Seitz[1,3]

Received: 5 February 2024 / Accepted: 7 May 2024

Abstract

In recent decades, perceptual learning (PL) has witnessed significant advancements, but the field has faced questions regarding reliability, replication issues, and challenges in understanding individual differences, hindering accessibility to diverse populations. To address these issues, we introduce PLFest, a novel, cross-platform UNITY-based app designed to promote accessible, reproducible, and open PL research. PLFest supports a variety of training and assessment procedures, focusing on psychophysics and PL of contrast sensitivity. It facilitates data collection on computers, tablets, and smartphones, enhancing accessibility and portability, making it ideal for large-scale, multi-site studies. PLFest aims to promote open science, data sharing, and reproducible research, fostering collaboration within the research community. As a first step, to validate PLFest, we conducted tests on healthy participants, assessing visual acuity and contrast sensitivity on both desktop and tablet setups. These measures were specifically chosen as they represent fundamental assessments of visual functions in healthy and clinical populations and are also known to be sensitive to display characteristics. The results demonstrated that PLFest produces reliable measurements, in particular on Apple iPad tablets, suggesting that the app is appropriate for visual psychophysics. This validation supports PLFest as a robust platform for PL research, emphasizing its potential to overcome limitations associated with high-end desktop/monitor setups and ensuring its applicability across diverse hardware configurations.

Keywords Perceptual learning · Open science platform · Reliability and validity · Visual acuity · Contrast sensitivity

Introduction

Perceptual learning (PL) refers to experienced-based changes, typically improvements, in the ability to extract sensory information from the environment and encompasses the set of mechanisms through which experience with the environment gives rise to changes in perceptual processing (Lu & Dosher, 2022; Sagi, 2011; Seitz, 2017). PL is fundamental to perceptual development, formation of perceptual expertise, and rehabilitation after sensory damage (Lu et al., 2016; Maniglia & Seitz, 2018; Seitz, 2017). From a scientific perspective, PL represents one of the most examined training phenomena, with numerous studies showing that a wide range of visual abilities can improve with practice. This includes processing stimulus orientation (Hung & Seitz, 2011; Jehee et al., 2012; Schoups et al., 1995; Schoups et al., 2001) and contrast (Adini et al., 2002; Deveau et al., 2014a, 2014b, 2014c; Furmanski et al., 2004; Polat et al., 2012), resolving fine detail (i.e., acuity; DeLoss et al., 2015; Deveau & Seitz, 2014; Deveau et al., 2014a, 2014b, 2014c; Polat et al., 2012), and higher-level visual abilities such as reading (Bernard et al., 2012; Chung et al., 2004; Deveau & Seitz, 2014; Lee et al., 2010; Polat et al., 2012; Yu et al., 2010). Beyond basic science, the field has generated numerous translational approaches (Deveau & Seitz, 2014; Deveau et al., 2013; Polat, 2009) aiming to exploit PL in interventions to enhance normal perceptual abilities (e.g., Athletes: (Appelbaum & Erickson, 2018; Deveau et al., 2014a, 2014b, 2014c), Medical Experts: (Kellman, 2013)) and to treat both purely perceptual (e.g., Amblyopia: (Hussain et al., 2014; Levi & Li, 2009; Li et al., 2008; Polat et al., 2009), Myopia: (Camilleri et al., 2014a, 2014b; Camilleri et al., 2014a, 2014b; Durrie & McMinn, 2007) Presbyopia: (Deveau et al.,

✉ Samyukta Jayakumar
samyukta.jayakumar@email.ucr.edu

[1] Department of Psychology, University of California, Riverside, Riverside, CA, USA

[2] Department of Psychology, University of Wisconsin Madison, Madison, WI, USA

[3] Department of Psychology, Northeastern University, Boston, MA, USA

Published online: 21 May 2024

2014a, 2014b, 2014c; Polat et al., 2012), Macular Degeneration: (Chung, 2011; Maniglia et al., 2020, 2016a, 2016b; Plank et al., 2014), and broader (e.g., ASD: (Harris et al., 2015), Dyslexia: (Gori et al., 2016)) disorders. The potential broader impacts of PL are immense and careful research in this domain can greatly enhance our basic understanding of the perceptual systems and the plasticity of these systems.

However, despite its substantial potential impact, a major obstacle to a clearer understanding of the mechanisms underlying PL, and, consequently, more successful translation, is that the field, to date, has been strongly driven by "novel" and "provocative" findings, often demonstrated via small N studies, with few research projects utilizing the type of large and heterogeneous samples that are necessary to achieve robust and unbiased results. In turn, unsurprisingly, the field of PL, like many others in psychology, has suffered from numerous replication challenges (Aberg & Herzog, 2010; Hung & Seitz, 2011, 2014; Liang et al., 2015a, 2015b; Xiao et al., 2008; Zhang et al., 2013; Zhang & Yu, 2016). This is undoubtedly exacerbated by the fact that PL researchers have largely examined the impact of their own paradigms via their own outcome measures, which frequently differ from those employed by other research groups in mostly unaccounted for ways. Such uncontrolled variability in approach severely hinders the field's ability to isolate the critical ingredients in successful PL (Hung & Seitz, 2011, 2014; Talluri et al., 2015).

As a glimpse into the inconsistencies within the existing PL literature, consider one of the more foundational results in the field that PL of orientation discrimination is specific to the location of training and involves plasticity in primary visual cortex (V1) (Schoups et al., 1995; Schoups et al., 2001). Although this initial result is still considered seminal in the field, subsequent physiological data from Ghose et al. (Ghose et al., 2002) failed to replicate the V1 finding, while other behavioral studies suggested that the observed degree of specificity is an artifact of pre-testing approaches. Together, these latter results appear to call the original findings into question. Yet, research by Seitz (Hung & Seitz, 2014) suggested that methodological differences, in particular in the adaptive procedures, could explain why later studies, such as (Zhang et al., 2010), did not show the same level of specificity observed in the original study. In the case of (Zhang et al., 2010), the staircases included more easy trials, which in turn led to less specificity. Modeling by Seitz (Sotiropoulos et al., 2018; Talluri et al., 2015) showed that this methodological difference is sufficient to explain the discrepancies in behavioral findings across groups. Additional work by Seitz utilizing deep neural networks (Wenliang & Seitz, 2018) has shown that differences in training thresholds are sufficient to account for the differences in neural findings found between Schoups (Schoups et al., 1995; Schoups et al., 2001) and Ghose (Ghose et al., 2002). Critically, many of these various methodological differences across studies, which appear to substantially alter the learning outcomes, were not purposeful manipulations. They instead reflect unintended differences in lab-specific methodological practices. These types of untracked and unaccounted for inconsistencies across studies abound in the literature.

It is a long-held field consensus that large-scale, multiple lab studies that include diverse individuals, regularize methods across labs, and use common account measures are necessary to advance rigor and reproducibility in the field. In 2008, the International Workshops on Perceptual Learning was formed (with the first meeting in Beijing) as a forum specifically for researchers in the field of perceptual learning. Here, the field established the goal of a PL "ModelFest" (with at least 10 additional PL ModelFest meetings held over the subsequent decade). This PL ModelFest (Klein et al., 2011) was originally meant to address the numerous surprising and minimally characterized findings in the field by replicating key PL results at multiple sites with conserved stimulus and task parameters and common outcome measures. These meetings identified a number of fundamental limitations of the field: (1) small sample sizes (often $N < 12$ per condition); (2) the use of reasonably homogenous participant samples that might not allow extrapolation to other populations, in particular those of translational interest (e.g., the use of young adult participants, who are frequently well-trained psychophysical observers, that might not be good models of the impact of PL on older adults); (3) lack of direct comparisons between training procedures; (4) lack of consistent generalization outcome measures; (5) lack of publicly available datasets upon which to advance models of PL; and (6) substantial hardware/software barriers to performing direct replications or extensions of existing work.

It was hoped that such a large-scale approach would clarify which methods show consistent results. However, a limitation of moving the project forward was a lack of appropriate tools to make the project convenient to participating labs. Indeed, labs can differ widely in terms of available hardware, software, and facilities, which in turn can make sharing and replication infeasible. In terms of hardware, for decades, vision science has relied on cathode ray tubes (CRTs), which allowed for highly accurate timing of stimulus presentation. However, with decreased production of these devices, researchers are turning towards using alternative equipment like liquid crystal displays (LCDs) for stimulus presentation (Rohr & Wagner, 2020; Zhang et al., 2018). While these are increasingly well-adopted, they exhibit different spatial–temporal display characteristics than CRTs. For example, while CRTs present briefly flashed points of light, LCDs present light continuously and then have brief periods of changes to luminance/color at frame transitions (Ghodrati et al., 2015). Because many display properties have the potential to fundamentally alter participant performance in myriad ways,

ensuring that results are reliable across hardware is a must. Then, in terms of software, tools like PsychToolBox and PsychoPy are most typical of PL studies. Although these tools are proven for lab-based research, they are challenging to share, and even more so to translate outside of the lab (especially to tablets and phones) as they require use of specific versions of MATLAB and Python (Nuutinen et al., 2018; Peirce et al., 2019). These hardware and software limitations pose a challenge when running studies across multiple sites, and even more so for replication.

Here, we introduce a tool to help tackle these challenges: an app called PLFest that is designed to promote accessible, reproducible, and open PL research. This app takes advantage of the tremendous development of consumer tablet technology where, for example, current iPad Pros support a 120-Hz screen refresh rate, over 300 pixels per inch, and 16 bits per color channel, making it a powerful psychophysical machine capable of ambulatory research and facilitating running many participants simultaneously. PLFest is innovative in a number of ways as it (1) makes available a number of training approaches that are representative of those that have been used in research in the field to-date, (2) includes numerous outcome measures (Fig. 1) by which to understand both near (e.g., to feature and task sets that have been trained) and far transfer effects (e.g., to features and task sets that are untrained, including a range of hearing, attentional and cognitive measures), (3) is highly configurable in a way that supports both easy reproduction of existing experiments and customization to support new experiments (including numerous built-in procedures and a scripting language to support more arbitrary procedures), and (4) is structured to be expandable to new methods (including both new task structures, stimulus sets, and use of peripherals such as eye-trackers and EEG systems). Importantly, with PLFest being built within the UNITY Game Engine, it is intrinsically cross-platform (it currently has been tested on iOS, Android, MacOS, and Windows), self-contained as a single app, and it does not rely upon versioning of other libraries and packages (unlike MATLAB and Python). Further, it supports easy configuration on multiple devices, including participant-owned hardware, and is integrated with a HIPAA compliant Amazon Web Services (AWS) back-end, supporting both lab-based and at-home-based research.

Data is saved in JSON file format or can be exported to CSVs, that can be easily read using several different programming languages such as MATLAB, Python, and R. A sampling of the range of tasks that PLFest supports is shown in Fig. 1. These include visual assessments (such as contrast sensitivity, contour detection, visual acuity, and search) and cognitive assessments (such as matrix reasoning, N-back, Cancellation, and Complex figure tasks). We note that there are many more tasks than the platform will allow, such as psychoacoustical tasks (pure tone thresholds, spectral, temporal and spectrotemporal sensitivity,

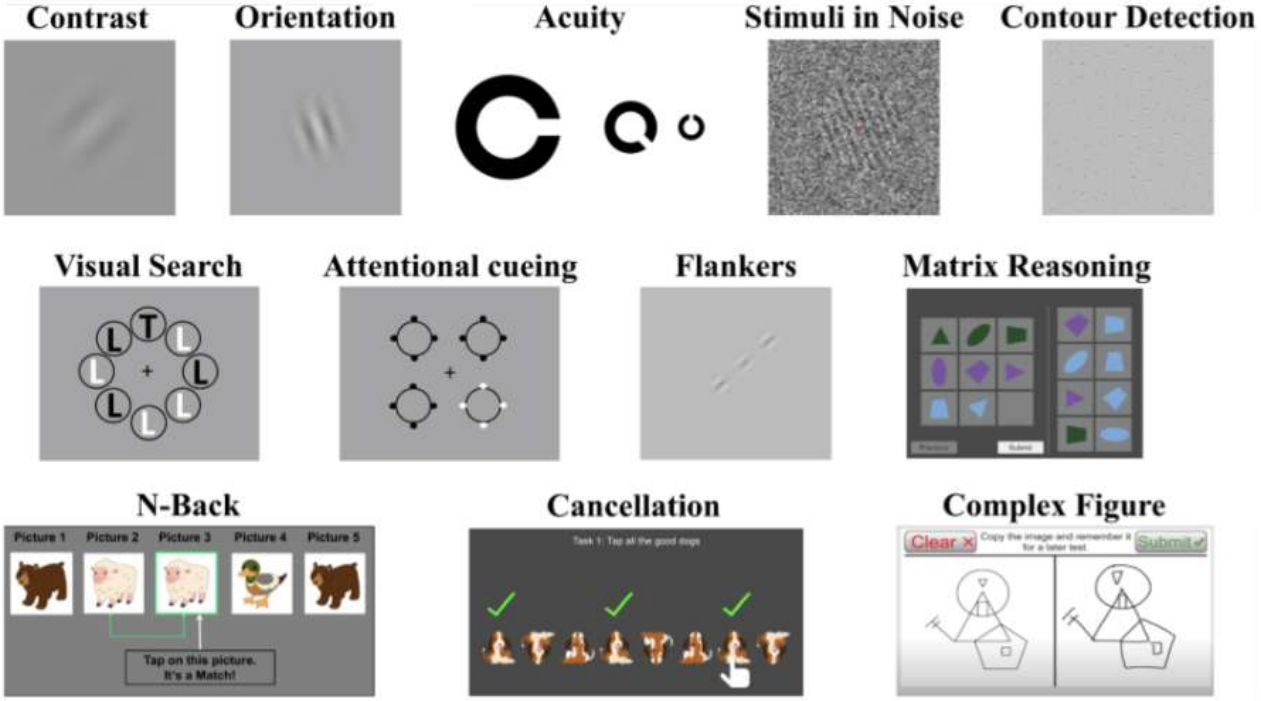

Fig. 1 Examples of tasks supported by the platform: In addition to basic visual tasks, the application can support a wide range of cognitive tasks such as matrix reasoning, n-back working memory task, cancellation, and complex figure tasks, among many others not displayed in the figure

binaural hearing, spatial release from masking, dichotic digits, etc.), and neuropsychological testing (such as word lists, constructional praxis, Boston naming, trail-making, and clock drawing), as well as reading and vocabulary tasks. A novelty of the platform is that parameters such as task durations/ trial numbers, stimulus durations, target sizes and locations, and adaptive procedures can be customized through the graphic user interface within the application, or by directly editing configuration files. The platform additionally supports the creation and implementation of questionnaires in a standalone format or within each task. It is also possible to combine multiple different tasks as separate batteries within a single experiment and include time-breaks, time-outs, and password locked sessions for flexible administration. Features that are still in development include eye-tracking, ambient sound monitoring, and communication with EEG systems. Furthermore, many other tasks, including those with moving stimuli, are under development. PLFest is currently in beta stage, and we encourage readers to contact the authors for instructions on how best to use the application.

As a first step towards validation of the PLFest platform, here, we examine reliability and consistency measurements of visual acuity (VA) and contrast sensitivity (CS), two of the most widely used tasks in PL. We tested these measures on a standard lab setup (monitor driven by a laptop) and a tablet setup. Additionally, we also measured the performance in the CS Task on a conventional display monitor (Cambridge Research Systems Display++) driven by a tablet and contrasted it with the performance on the tablet setup. VA measures the ability to distinguish fine structure of objects, usually at high contrasts, and is considered a fundamental measure of spatial vision. CS meanwhile measures the ability to identify differences between the luminance of the object and its background and is considered a fundamental measure of light sensitivity. VA and CS are foundational in evaluating visual performance in clinical and experimental settings and can be predictors of visual pathologies such as cataracts, glaucoma, diabetic retinopathy, and macular degeneration (Brown et al., 2002; Howes et al., 1982; Klein et al., 1995; Lin et al., 2018; Ross et al., 1984; Sabour-Pickett et al., 2013). Critically, both tests are also particularly sensitive to display parameters, with visual acuity requiring high-pixel density and spatial precision, and contrast sensitivity requiring linearization display gammas and precision of luminance values. As we will see below, results showed equal-to-greater reliability of PLFest both on tests of acuity and contrast sensitivity on an iPad Pro compared to LCD display as well as good inter-device reliability of CS measures on all three displays. These data support the use of PLFest as a platform for perceptual learning research.

Methods

Participants

Fifty undergraduate students (16 M, 1 other; mean age = 19.1 years [SD = 1.16]) at the University of California, Riverside (UCR) participated in the study comparing the tablet and a standard lab setup (LG monitor). For measuring the consistency across the tablet and Display++ devices we additionally recruited 48 undergraduate students (21 M; Mean age = 20.7 years [SD = 4.6]) at UCR. Written and informed consent was obtained from all participants in the study and all participants were given research credits for participation. All participants reported normal or corrected to normal vision ($\geq$ 20/40), and the study was approved by the IRB of UC Riverside.

Outlier Exclusion

Thresholds were calculated by taking the average of the last 6 reversal values of participant performance in the session phase for each task and on each device. We employed a 2-step outlier exclusion for the data collected in the study. The first step involved excluding participants whose thresholds were above 10% contrast (> -1 logCS) and 0.3 logMAR (translating to a Snellen acuity of 20/40), respectively. This was performed to exclude thresholds that were more likely due to attentional lapses or general issues with task comprehension and does not reflect task performance. The second step excluded participants based on the distance with respect to the dispersion of mean, i.e., thresholds $\pm$ 3 SD of the mean were excluded. The total *N* for the two tasks after exclusions were 41 (for CS task) participants and 48 (for VA task) in the study comparing performance between the tablet and LG monitor. Three participants were excluded from the analyses comparing performance of participants between the tablet and Display++ device for the CS task.

Setup

Participants in the condition comparing the tablet and a standard lab setup performed the tasks on a tablet (2021 Apple iPad pro-12.9″ with a resolution of 2048 × 2732 and refresh rate of 120 Hz) and a monitor (LG monitor with a resolution of 1920 × 1080 and refresh rate of 120 Hz) driven by a separate laptop (Alienware, Dell Inc.). For participants in the condition contrasting a more conventional display device with the tablet, we measured performance on a Cambridge Research Systems Display++ that was driven directly by the 2021 Apple iPad Pro. Participants performed under low-light conditions, while the background luminance

of both devices was matched (92.5 cd/m^2, background luminance of Display++ and iPad was matched at 74.7 cd/m^2). Head and chin rests were used with the devices to minimize discrepancies in viewing distance across participants. The tablet was mounted on a stand and placed at a viewing distance of 50 cm; the LG monitor was placed at a distance of 67 cm and the Display++ screen at 91 cm in order to maintain equal size of the stimulus in visual angle across all devices. Participants were provided with headphones for auditory feedback, and they recorded their responses using the arrow keys on a keypad.

Procedure

Participants performed a 1-h session where they were tested on two tasks: (1) CS and (2) VA. They performed two runs of each task on two devices: monitor (LG monitor or Display++) and tablet. Prior to performing the full session, participants were briefly exposed to each task on both devices in order to ensure they understood the tasks and knew what the stimuli would look like on both devices. The overall study design is shown in Fig. 2.

The order of devices was counterbalanced across participants. Participants always performed the CS task followed by the VA task during the practice and in each of the two runs.

Tasks

The details of the two tasks used in the study are as follows.

CS Task

This test uses centrally presented Gabor patches with a spatial frequency of 6 cpd tilted at either 45° or 135° to estimate contrast sensitivity (Fig. 3A). Participants were required to indicate the direction of tilt of the Gabor using the left/right arrow keys on the keypad. Participants first performed a 2-stage practice of this task on each device where the first stage consisted of a total of 10 trials during which the stimulus was presented with progressive difficulty starting at a value of -0.4 logCS with a step factor of 0.3 log units. This was followed by the second stage that consisted of a conventional 2-down 1-up staircase algorithm terminating

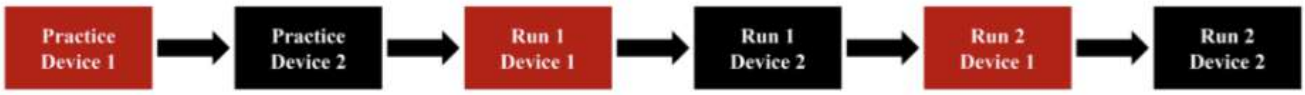

Fig. 2 Overview of study design: The red and black boxes denote specific devices (i.e., monitor and tablet), based on the order of each device presented to the participant). Prior to beginning the experiment, participants were required to practice both the tasks on each of the devices after which they completed Run 1 on each device before completing Run 2 on the devices. The order of the tasks performed was kept constant (i.e., CS followed by VA) throughout the study

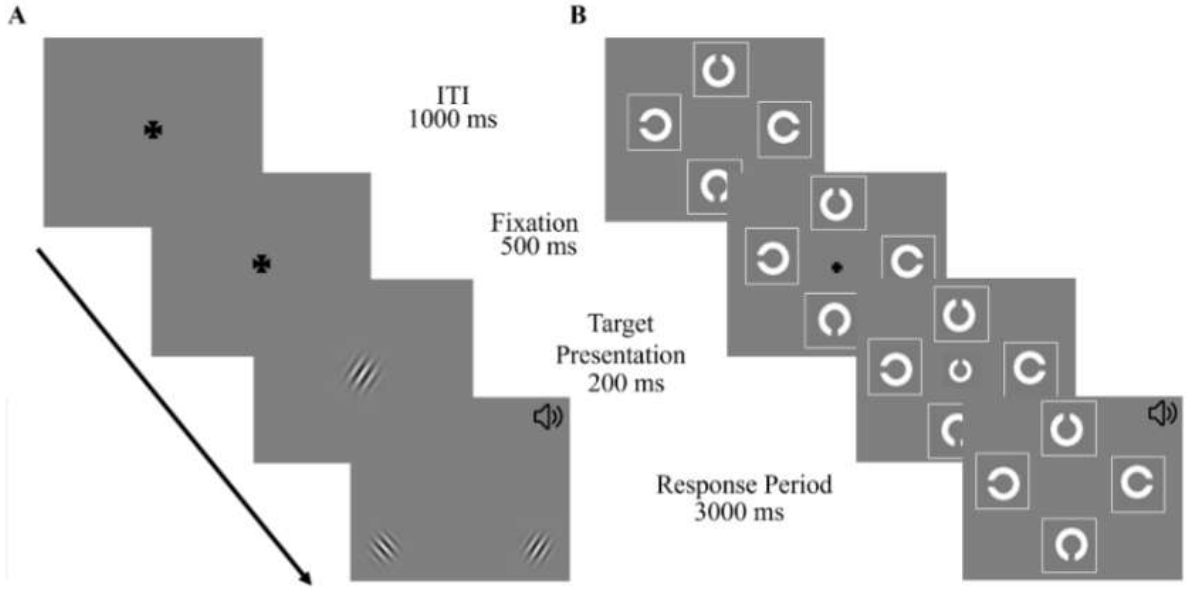

Fig. 3 CS (**A**) and VA (**B**) task paradigms: Participants respond to the target orientation using the arrow keys on a keypad and receive auditory feedback based on the accuracy of their response, i.e., high pitch pleasant tone for correct and a low beep for incorrect responses

after 20 trials, starting at a value of -0.1 logCS and a step size of 0.3 log units. Once practice was completed on both the devices, participants performed a longer version of this task. In order to mitigate the potential effects of cold-start performance as well as device and task switching effects, participants first performed the two-stage practice as mentioned above prior to engaging in the main task. The main task employed a 2-stage conventional 3-down/1-up staircase with a step factor of 0.3 log units until 3 reversals (stage 1), after which the step factor remained 0.1 log units (stage 2) until task completion.

When measuring the consistency between the iPad and Display ++ devices, the general structure of the contrast task was the same as mentioned above. The only difference is that a blockwise staircase was used for the main task where 10 mini blocks of 6 trials were implemented with each of the 6 trials corresponding to 6 different orientations of the Gabor (22.5°, 45°, 67.5°, 112.5°, 135°, and 157.5°) in random order. Contrast of the Gabors stepped down (i.e., more difficult) when there were 1 or fewer errors within the block, stayed the same if there were 2 errors, and stepped up (i.e., easier) when 3 or more errors were made within each block. This allowed for maintaining the performance between 66% and 83% as an average across the orientations. In this case, data was averaged across the blocks to ensure that there were adequate numbers of steps used to obtain a threshold. Manipulations to the CS task structure were made such that it is consistent with a larger clinical study that is currently ongoing in the lab. Of note, data depicting contrast thresholds in both the tablet and Display ++ devices shown in Fig. 6 was collected using this structure of the CS task.

VA Task

In this test, a block letter C (Sloan Font) is presented in one of four orientations (with the gap of the C facing up, down, right or left) as shown in Fig. 3B. Participants were required to indicate the side of the gap using the up, down, right/left arrow keys on the keypad. The structure of this task was slightly different from the CS task with two main stages. In alignment with the CS task, participants first performed a 2-stage practice on both the devices where the first stage comprised a total of 12 trials with progressive difficulty (1.3, 1.1, 0.9, 0.7, 0.5, 0.4, 0.3, 0.2, 0.1, 0.0, − 0.1, and − 0.2 logMAR units) followed by a conventional 2- down 1-up staircase in the second stage with a starting value of 0.5 logMAR and a step factor of 0.2 log units that terminated after 20 trials. Post the completion of practice on both the devices, participants performed the first stage of the practice followed by the main task which employed a conventional 3-down 1-up staircase that terminated after 60 trials. Similar to the CS task, here, the main task also consisted of 2 stages where we used a step factor of 0.2 log units until 3 reversals (stage 1) after which the step factor remained at 0.1 log units (stage 2) until completion.

Results

We first examine within-device reliability (i.e., repeatability on the same device), followed by inter-device reliability (i.e., consistency of the measurements across platforms) on each of the tasks respectively.

Within-Device Reliability

The first question we asked was whether PLFest leads to reliable, repeatable measures within each device on which it is run.

Contrast Sensitivity

Within-device reliability of the CS task on the monitor (Fig. 4A) and the tablet (Fig. 4B), respectively, showed correlations of moderate magnitude for both devices $r_{\text{Monitor}}(39) = 0.72$, $p_{\text{Monitor}} < 0.001$; $r_{\text{Tablet}}(39) = 0.6$, $p_{\text{Tablet}} < 0.001$. To understand these relationships in more detail, we examined Bland–Altman's limits of agreement (LoA), which evaluates the difference in thresholds (Run 2–Run 1) between the two runs as a function of the average thresholds (mean of test–retest) across the two runs (Altman & Bland, 2017; Bland & Altman, 1999). To illustrate the between-subject variability of threshold estimation, the mean across runs is shown on the *x*-axis to provide a single point estimate for each participant in terms of their overall estimated threshold. To display the within-subject variability of the predicted threshold, the difference between the two runs is plotted on the *y*-axis to provide a single point estimate of the extent of divergence of performance between the two runs. The main point estimate of the systematic bias in the measurement across sessions is represented by the distance between the mean of these discrepancies, which is depicted as a straight line perpendicular to the *y*-axis and symbolizes zero (zero = perfect agreement). Plotted as dotted lines, the 95% LoA [± 1.96 SD (difference between sessions)] represents an estimate of the region in which 95% of the within-subject, between-run changes of threshold estimates are likely to be seen. As observed from Fig. 4C and D, the mean difference between the runs was close to 0 in all of the within-device comparisons indicative of less systematic bias within the two devices. Overall, the data shows that CS tests were reliable on both platforms.

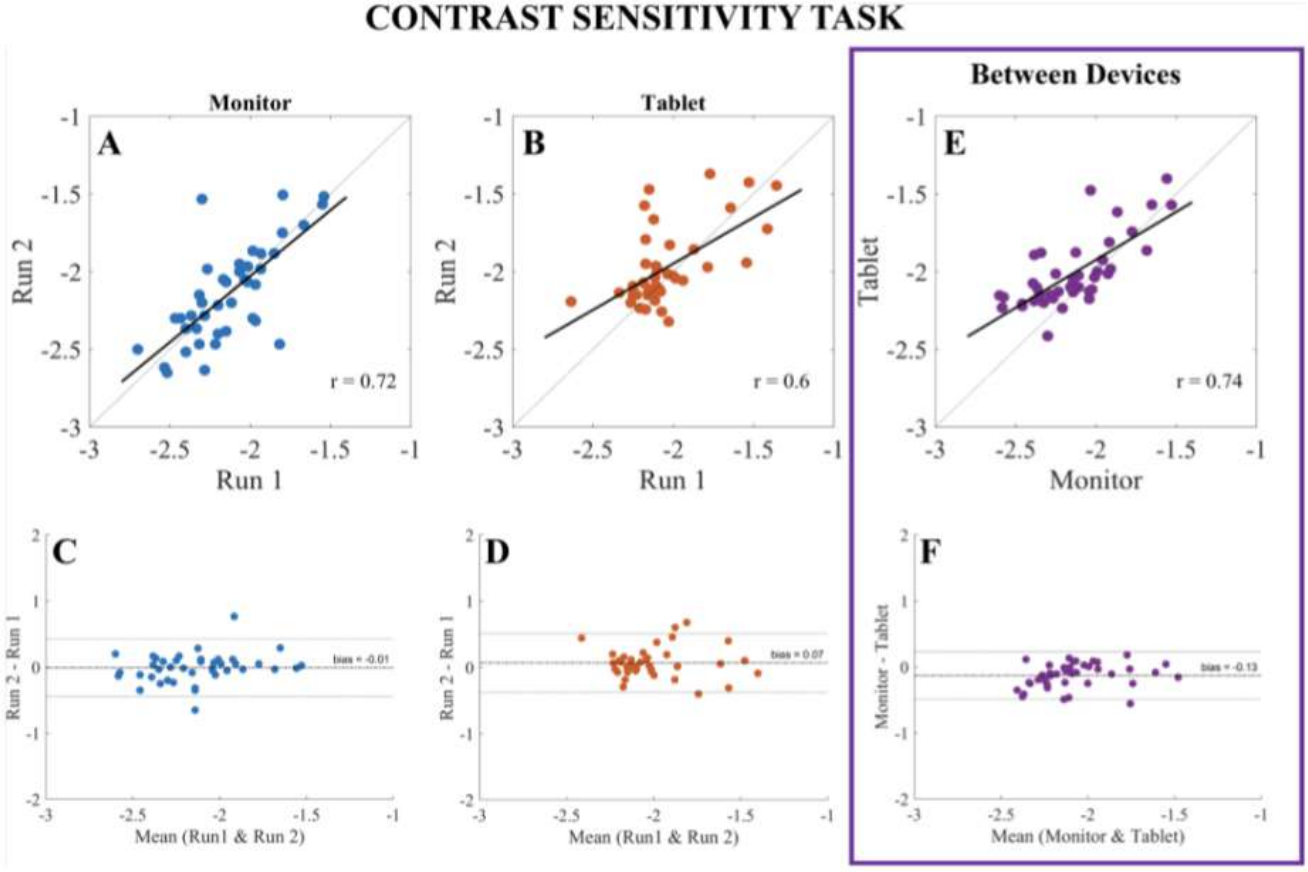

Fig. 4 Performance on CS Task within (**A–D**) and between (**E, F**) devices: Participant thresholds in run 1 (*x*-axis) are plotted against their thresholds in run 2 (*y*-axis) on the monitor (**A**) and tablet (**B**) devices, respectively. Bland–Altman LoA for within-device thresholds on the monitor (**C**) and tablet (**D**) tasks. The mean threshold across both runs (*x*-axis) is plotted against the difference between the two runs (*y*-axis). Solid line denotes the mean difference between runs and the dotted lines indicate 95% LoA. A negative value on the *y*-axis denotes better performance on the second run for both tasks. Between device correlations and LoA plots can be observed in the right most panel (**E, F**), respectively. Here participant thresholds were averaged across both runs for each device and performance was correlated between the two devices (**E**). Bland–Altman LoA for between-device comparisons (**F**) shows the mean threshold across both devices (*x*-axis) plotted against the difference between the two devices (*y*-axis). Solid line denotes the mean difference between the devices and the dotted lines indicate 95% LoA. Here, a positive value on the *y*-axis denotes better performance on the tablet device

Visual Acuity

Within-device reliability of the VA task on the monitor (Fig. 5A) and the tablet (Fig. 5B), respectively, showed a correlation of moderate magnitude in the tablet ($r_{\text{Tablet}}(46) = 0.7$, $p_{\text{Tablet}} < 0.001$), but poor reliability on the monitor for this task ($r_{\text{Monitor}}(46) = 0.24$, $p_{\text{Monitor}} = 0.1$). A closer examination of the LoA plots for test–retest reliability on the monitor (Fig. 5C) reveal a more dispersed distribution of differences in performance between the two runs with a low mean difference (bias = 0.01) whereas a more clustered distribution for the tablet (Fig. 5D) device with a mean difference of 0.02. A possible explanation for this may be the lower resolution of the monitor compared to the tablet; however, these data show that PLFest is producing reliable data on the tablet platform for visual acuity measures.

Inter-device Reliability

We next asked whether PLFest, when running on a tablet, assesses the same constructs as a traditional vision science setup using a desktop monitor. To examine this, we measured inter-device reliability (e.g., correlations between participants' performance across the two devices) for CS (Fig. 4E) and VA (Fig. 5E). In this analysis, we averaged the thresholds across both runs for each participant and task. We observed moderate correlations for both CS ($r_{\text{CS}} = 0.74$, $p < 0.001$) and VA ($r_{\text{VA}} = 0.55$, $p < 0.001$) tasks between the two devices that are within range of the within-device reliabilities for each test. Figures 4F and 5F show the Bland–Altman plots, for the CS and VA tasks, respectively, to test for the consistency in measuring the performance on both tasks between the two devices. To do this, we averaged the thresholds across both runs for each participant and task.

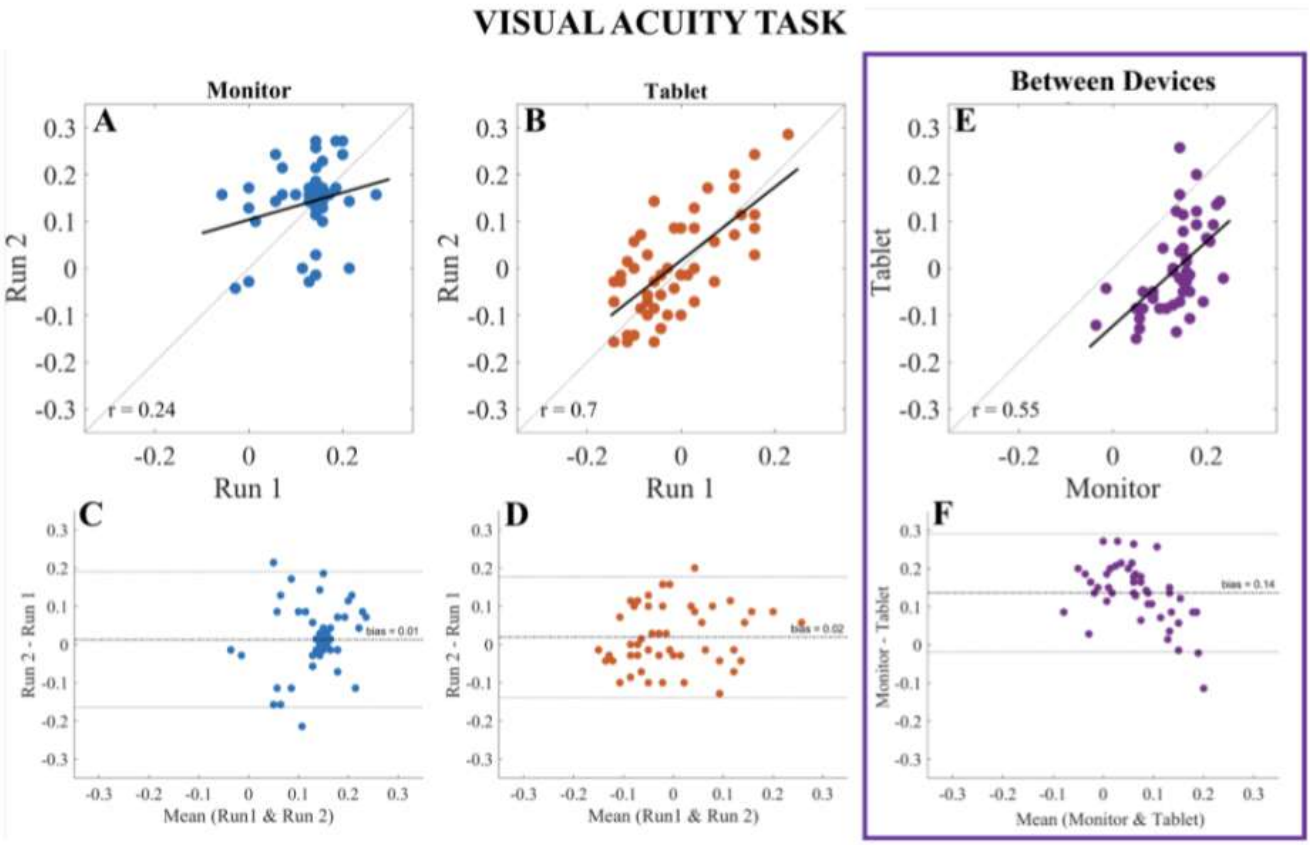

Fig. 5 Performance on VA Task within (**A–D**) and between (**E, F**) devices: Participant thresholds in run 1 (*x*-axis) are plotted against their thresholds in run 2 (*y*-axis) on the monitor (**A**) and tablet (**B**) devices respectively. Bland–Altman LoA for within-device thresholds on the monitor (**C**) and tablet (**D**) tasks. The mean threshold across both runs (*x*-axis) is plotted against the difference between the two runs (*y*-axis). Solid line denotes the mean difference between runs and the dotted lines indicate 95% LoA. A negative value on the *y*-axis denotes better performance on the second run for both tasks. Between device correlations and LoA plots can be observed in the right most panel (**E, F**) respectively. Here participant thresholds were averaged across both runs for each device and performance was correlated between the two devices (**E**). Bland–Altman LoA for between-device comparisons (**F**) shows the mean threshold across both devices (*x*-axis) plotted against the difference between the two devices (*y*-axis). Solid line denotes the mean difference between the devices and the dotted lines indicate 95% LoA. Here, a positive value on the *y*-axis denotes better performance on the tablet device

The *x*-axis shows the average thresholds of the participants across both the devices and the *y*-axis represents the difference in thresholds between the monitor and tablet. A positive bias in both the CS and VA task is indicative of better performance on the tablet in comparison to the monitor. From these graphs it can be observed that participants perform better on the tablet when compared to the monitor on the VA task in contrast to the CS task ($bias_{CS} = -0.13$, $bias_{VA} = 0.14$, although the bias is small in both cases), perhaps due to the greater spatial and luminance resolution on the tablet compared to the desktop display.

To further test whether running PLFest on a tablet leads to reliable inter-device performance, we compared PLFest on the tablet with a more conventional Display++ monitor used in vision science, and one that is specialized to produce high contrast depth. To do this, we tested a modified version of the CS task as indicated in the methods. Figure 6A shows the moderate and significant correlations ($r_{CS} = 0.6$, $p < 0.001$) between the thresholds on the tablet and the Display++ devices. It can also be observed from the Bland–Altman plots (Fig. 6B) as indicated by the small positive bias ($bias_{CS} = 0.02$) that participants performed slightly better on the tablet compared to the Display++. These results suggest that the iPad can produce comparative data to the research grade Display++ system.

Discussion

The goal of this study was to introduce and validate PLFest, a novel, cross-platform app to support perceptual learning research. Here, we found that measures of contrast sensitivity and visual acuity show acceptable within-device reliability when tested on an iPad tablet. Further tests of inter-device reliability show that the iPad produces results that are consistent with those found on consumer grade LCD

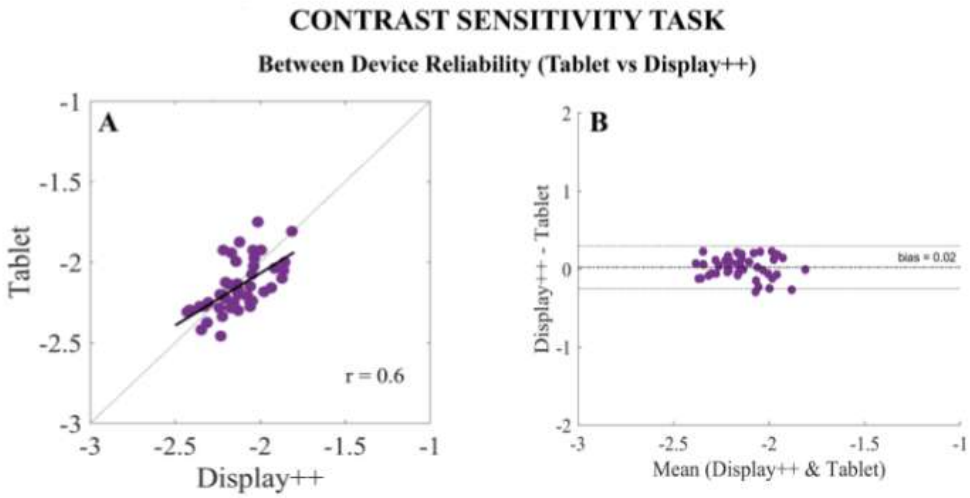

Fig. 6 Performance on CS Task between the tablet and Display++ devices: Between-device correlations and LoA plots can be observed in panels **A** and **B** respectively. Here participant thresholds were averaged across both runs for each device and performance was correlated between the two devices (**A**). Bland–Altman LoA for between-device comparisons (**B**) shows the mean threshold across both devices (*x*-axis) plotted against the difference between the two devices (*y*-axis). The solid line denotes the mean difference between the devices and the dotted lines indicate 95% LoA. Here, a positive value on the *y*-axis denotes better performance on the tablet device

and the research grade Display++ system (although both systems have lower spatial resolution than the iPad). These results show that PLFest is a reliable platform for visual psychophysics and can support PL research.

Over the last several years, multiple tests measuring CS and VA have been developed on multiple platforms such as computer devices and made commercially available for testing. However, it was only until recently that these tests were created for use on remote devices such as iPads and tablets (Habtamu et al., 2019; Kollbaum et al., 2014; Labiris et al., 2023). For example, (Habtamu et al., 2019) developed and validated a smartphone based CS test employing the tumbling E Pelli-Robson CS test and PeekCS test on Android phones and observed high test–retest correlations. Similarly, (Kollbaum et al., 2014) validated an iPad-based Letter CS test and compared it to the Pelli-Robson test and Freiburg Acuity and Contrast Test (FrACT) and observed good repeatability. On the other hand, (Labiris et al., 2023) recently validated a web based Democritus Digital Acuity and Reading test (DDART) against conventionally used distance vision charts across multiple sites with high n (543 participants) and noticed reliable measurements in both normal and low vision patients. While all of these studies show reliable measures akin to ours, our platform offers a one-stop solution to conduct multiple vision-based tests on a single platform. Further, it is structured to facilitate both testing and perceptual learning training within the same platform.

The development of PLFest builds upon other developments cropping out from the COVID-19 pandemic emphasizing the need for reliable and remote testing platforms worldwide (Collins et al., 2022; Yaghoubi et al., 2022) and telehealth approaches more generally. This has led our research group to develop several tools to measure several different auditory (Lelo de Larrea-Mancera et al., 2020) and cognitive functions (Pahor et al., 2019, 2022) on tablet-based devices that are reliable and portable to reach more diverse populations and adequately support comparisons across different approaches developed by different research groups. PLFest is built within the same test framework as these other tasks and while in the current paper we emphasized measurements of basic visual functions, the app is capable of doing more than just measuring VA and CS. Indeed, the current version of PLFest includes visual search tasks, orientation discrimination, contour integration, numerous psychoacoustic tasks, cognitive and neuropsychological tasks, for visual, auditory perceptual learning and cognitive training.

While these results are encouraging, there are also some limitations to the current study. First, we notice that a few participants were outliers (i.e., had unreliable scores/contrast thresholds above 10%). This is most likely due to issues with task instructions or lapses of attention. This is not uncommon in psychophysical research with undergraduate students; however, it emphasizes the importance of uniform instructions across participants, sufficient breaks between tasks and potentially employing performance cutoffs during practice that could be informative of, and further mitigate such occurrences. These issues are being addressed in versions of PLFest currently being developed, specifically via the implementation of animated tutorials and automatic checks for outlying performance. Closer examination of the

data from the VA task data suggests that the tablet might be more reliable than the monitor, which is likely a consequence of the lower poor resolution on the monitor screen compared to the iPad. This could have led to the pixel density not being high enough to capture an acuity better than 20/25, thus contributing to poor VA thresholds when measured on this device. Future use of PLFest in combination with desktop monitors should use displays of higher resolution (or have participants at a further distance).

The PLFest app is currently validated on iPad devices, due to their high resolution and refresh rate, allowing for high quality rendering of the stimuli. In terms of field of view, at the distance (50 cm) it was tested in the current study, the screen subtended approximately 32° × 24° of visual angle, and it can be placed as close as 35 cm (subtending approximately 46° × 34° of visual angle) without sacrificing conform and usability. Further, the app is cross-platform and can also run on Mac and PC Desktops and also other iOS and Android tablets and even phones, and we plan future studies to validate across a larger range of platforms. While the platform currently does not include visual assessments with moving stimuli, these are planned for future studies, and given the high spatial and temporal resolution of tablets and phones (catering users that expect high-fidelity movies), we are confident that psychophysics related to visual motion can be conducted with psychophysical precision on mobile platforms. Further, with remote use in mind, next steps will include moving beyond the present study's in-lab setting and assess reliability and consistency of PLFest in remote testing conditions that are susceptible to several uncontrolled environmental factors such as lighting, viewing distance, screen brightness etc. We note that this has already be accomplished in the case of psychoacoustics (Lelo de Larrea-Mancera et al., 2022), which is also quite sensitive to the perceptual environment in which testing is conducted.

Overall, PLFest shows great promise as a reliable cross-platform tool to promote open research in perceptual learning. As it is publicly available and a free to use platform, PLFest opens up a multitude of opportunities to conduct a wide range of vision science experiments both in-in-lab as well as in remote settings, which can facilitate research in underserved communities. Further the platform supports straightforward localization to different languages. PLFest is structured both to facilitate easy reproducibility of perceptual learning research, as well as comparisons across studies through the use of common outcome measures. Further the platform can support modeling both through its standardization of data structures as well as through APIs to directly interact with models that are currently under development. Overall, PLFest can open up new possibilities for studying vision and perceptual learning and help address the long-standing issues of replicability and reproducibility in the field.

Author Contribution All authors contributed to the study conception and design. Data collection, analysis, and interpretation, as well as manuscript preparation were performed by SJ and MM. Data interpretation and manuscript preparation were performed by CSG and ARS. Additionally, both CSG and ARS secured funding for this study from the National Eye Institute (NEI R01 EY031226 grant).

Funding This study was supported by the National Eye Institute NEI R01 EY031226 grant.

Data Availability Data will be made available upon request. Please direct all requests to the corresponding author at samyukta.jayakumar@email.ucr.edu.

Declarations

Ethics Approval This study has been approved by the University of California, Riverside, Institutional Review Board (IRB). All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional research committee and with the 1964 Helsinki declaration and its later amendments.

Conflict of Interest Authors Seitz and Green have editorial roles with the Journal of Cognitive Enhancement.

References

Aberg, K. C., & Herzog, M. H. (2010). Does perceptual learning suffer from retrograde interference? *PLoS ONE, 5*(12), e14161. https://doi.org/10.1371/journal.pone.0014161

Adini, Y., Sagi, D., & Tsodyks, M. (2002). Context-enabled learning in the human visual system. *Nature, 415*(6873), 790–793. https://doi.org/10.1038/415790a

Altman, D. G., & Bland, J. M. (2017). Assessing agreement between methods of measurement. *Clinical Chemistry, 63*(10), 1653–1654. https://doi.org/10.1373/clinchem.2016.268870

Appelbaum, L. G., & Erickson, G. (2018). Sports vision training: A review of the state-of-the-art in digital training techniques. *International Review of Sport and Exercise Psychology, 11*(1), 160–189. https://doi.org/10.1080/1750984X.2016.1266376

Bernard, J.-B., Arunkumar, A., & Chung, S. T. L. (2012). Can reading-specific training stimuli improve the effect of perceptual learning on peripheral reading speed? *Vision Research, 66*, 17–25. https://doi.org/10.1016/j.visres.2012.06.012

Bland, J. M., & Altman, D. G. (1999). Measuring agreement in method comparison studies. *Statistical Methods in Medical Research, 8*(2), 135–160. https://doi.org/10.1177/096228029900800204

Brown, M. M., Brown, G. C., Sharma, S., Landy, J., & Bakal, J. (2002). Quality of life with visual acuity loss from diabetic retinopathy and age-related macular degeneration. *Archives of Ophthalmology, 120*(4), 481–484. https://doi.org/10.1001/archopht.120.4.481

Camilleri, R., Pavan, A., Ghin, F., Battaglini, L., & Campana, G. (2014a). Improvement of uncorrected visual acuity and contrast sensitivity with perceptual learning and transcranial random noise stimulation in individuals with mild myopia. *Frontiers in Psychology, 5*, 1234. https://doi.org/10.3389/fpsyg.2014.01234

Camilleri, R., Pavan, A., Ghin, F., & Campana, G. (2014b). Improving myopia via perceptual learning: Is training with lateral masking the only (or the most) efficacious technique? *Attention, Perception & Psychophysics, 76*(8), 2485–2494. https://doi.org/10.3758/s13414-014-0738-8

Chung, S. T. L. (2011). Improving reading speed for people with central vision loss through perceptual learning. *Investigative*

Ophthalmology & Visual Science, 52(2), 1164–1170. https://doi.org/10.1167/iovs.10-6034

Chung, S. T. L., Legge, G. E., & Cheung, S.-H. (2004). Letter-recognition and reading speed in peripheral vision benefit from perceptual learning. *Vision Research, 44*(7), 695–709. https://doi.org/10.1016/j.visres.2003.09.028

Collins, C. L., Pina, A., Carrillo, A., Ghil, E., Smith-Peirce, R. N., Gomez, M., Okolo, P., Chen, Y., Pahor, A., Jaeggi, S. M., & Seitz, A. R. (2022). Video-based remote administration of cognitive assessments and interventions: A comparison with in-lab administration. *Journal of Cognitive Enhancement, 6*(3), 316–326. https://doi.org/10.1007/s41465-022-00240-z

DeLoss, D. J., Watanabe, T., & Andersen, G. J. (2015). Improving vision among older adults: Behavioral training to improve sight. *Psychological Science, 26*(4), 456–466. https://doi.org/10.1177/0956797614567510

Deveau, J., Lovcik, G., & Seitz, A. R. (2013). The therapeutic benefits of perceptual learning. *Current Trends in Neurology*, *7*, 39–49. https://www.ncbi.nlm.nih.gov/pubmed/25580062. Accessed 20 Dec 2023

Deveau, J., Jaeggi, S. M., Zordan, V., Phung, C., & Seitz, A. R. (2014a). How to build better memory training games. *Frontiers in Systems Neuroscience, 8*, 243. https://doi.org/10.3389/fnsys.2014.00243

Deveau, J., Lovcik, G., & Seitz, A. R. (2014b). Broad-based visual benefits from training with an integrated perceptual-learning video game. *Vision Research, 99*, 134–140. https://doi.org/10.1016/j.visres.2013.12.015

Deveau, J., Ozer, D. J., & Seitz, A. R. (2014c). Improved vision and on-field performance in baseball through perceptual learning. *Current Biology: CB, 24*(4), R146–R147. https://doi.org/10.1016/j.cub.2014.01.004

Deveau, J., & Seitz, A. R. (2014). Applying perceptual learning to achieve practical changes in vision. *Frontiers in Psychology, 5*, 1166. https://doi.org/10.3389/fpsyg.2014.01166

Durrie, D., & McMinn, P. S. (2007). Computer-based primary visual cortex training for treatment of low myopia and early presbyopia. *Transactions of the American Ophthalmological Society*, *105*, 132–138; discussion 138–140. https://www.ncbi.nlm.nih.gov/pubmed/18427602. Accessed 20 Dec 2023

Furmanski, C. S., Schluppeck, D., & Engel, S. A. (2004). Learning strengthens the response of primary visual cortex to simple patterns. *Current Biology: CB, 14*(7), 573–578. https://doi.org/10.1016/j.cub.2004.03.032

Ghodrati, M., Morris, A. P., & Price, N. S. C. (2015). The (un)suitability of modern liquid crystal displays (LCDs) for vision research. *Frontiers in Psychology, 6*, 303. https://doi.org/10.3389/fpsyg.2015.00303

Ghose, G. M., Yang, T., & Maunsell, J. H. R. (2002). Physiological correlates of perceptual learning in monkey V1 and V2. *Journal of Neurophysiology, 87*(4), 1867–1888. https://doi.org/10.1152/jn.00690.2001

Gori, S., Seitz, A. R., Ronconi, L., Franceschini, S., & Facoetti, A. (2016). Multiple causal links between magnocellular–dorsal pathway deficit and developmental dyslexia. *Cerebral Cortex, 26*(11), 4356–4369. https://doi.org/10.1093/cercor/bhv206

Habtamu, E., Bastawrous, A., Bolster, N. M., Tadesse, Z., Callahan, E. K., Gashaw, B., Macleod, D., & Burton, M. J. (2019). Development and validation of a smartphone-based contrast sensitivity test. *Translational Vision Science & Technology, 8*(5), 13. https://doi.org/10.1167/tvst.8.5.13

Harris, H., Israeli, D., Minshew, N., Bonneh, Y., Heeger, D. J., Behrmann, M., & Sagi, D. (2015). Perceptual learning in autism: Over-specificity and possible remedies. *Nature Neuroscience, 18*(11), 1574–1576. https://doi.org/10.1038/nn.4129

Howes, S. C., Caelli, T., & Mitchell, P. (1982). Contrast sensitivity in diabetics with retinopathy and cataract. *Australian Journal of Ophthalmology, 10*(3), 173–178. https://doi.org/10.1111/j.1442-9071.1982.tb00380.x

Hung, S.-C., & Seitz, A. R. (2011). Retrograde interference in perceptual learning of a peripheral hyperacuity task. *PLoS ONE, 6*(9), e24556. https://doi.org/10.1371/journal.pone.0024556

Hung, S.-C., & Seitz, A. R. (2014). Prolonged training at threshold promotes robust retinotopic specificity in perceptual learning. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience, 34*(25), 8423–8431. https://doi.org/10.1523/JNEUROSCI.0745-14.2014

Hussain, Z., Astle, A. T., Webb, B. S., & McGraw, P. V. (2014). The challenges of developing a contrast-based video game for treatment of amblyopia. *Frontiers in Psychology, 5*, 1210. https://doi.org/10.3389/fpsyg.2014.01210

Jehee, J. F. M., Ling, S., Swisher, J. D., van Bergen, R. S., & Tong, F. (2012). Perceptual learning selectively refines orientation representations in early visual cortex. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience, 32*(47), 16747–16753a. https://doi.org/10.1523/JNEUROSCI.6112-11.2012

Kellman, P. J. (2013). Adaptive and perceptual learning technologies in medical education and training. *Military Medicine, 178*(10 Suppl), 98–106. https://doi.org/10.7205/MILMED-D-13-00218

Klein, R., Wang, Q., Klein, B. E., Moss, S. E., & Meuer, S. M. (1995). The relationship of age-related maculopathy, cataract, and glaucoma to visual acuity. *Investigative Ophthalmology & Visual Science*, *36*(1), 182–191. https://www.ncbi.nlm.nih.gov/pubmed/7822146. Accessed 11 Dec 2023

Klein, S., Carney, T., Yu, C., & Levi, D. (2011). The modelfest perceptual learning initiative: A status report. *Journal of Vision, 11*(11), 976–976. https://doi.org/10.1167/11.11.976

Kollbaum, P. S., Jansen, M. E., Kollbaum, E. J., & Bullimore, M. A. (2014). Validation of an iPad test of letter contrast sensitivity. *Optometry and Vision Science: Official Publication of the American Academy of Optometry, 91*(3), 291–296. https://doi.org/10.1097/OPX.0000000000000158

Labiris, G., Panagiotopoulou, E.-K., Delibasis, K., Duzha, E., Bakirtzis, M., Panagis, C., Boboridis, K., Mokka, A., Balidis, M., Damtsi, C., & Ntonti, P. (2023). Validation of a web-based distance visual acuity test. *Journal of Cataract and Refractive Surgery, 49*(7), 666–671. https://doi.org/10.1097/j.jcrs.0000000000001176

Lee, H.-W., Kwon, M., Legge, G. E., & Gefroh, J. J. (2010). Training improves reading speed in peripheral vision: Is it due to attention? *Journal of Vision, 10*(6), 18. https://doi.org/10.1167/10.6.18

Lelo de Larrea-Mancera, E. S., Stavropoulos, T., Carrillo, A. A., Cheung, S., He, Y. J., Eddins, D. A., Molis, M. R., Gallun, F. J., & Seitz, A. R. (2022). Remote auditory assessment using Portable Automated Rapid Testing (PART) and participant-owned devices. *The Journal of the Acoustical Society of America, 152*(2), 807. https://doi.org/10.1121/10.0013221

Lelo de Larrea-Mancera, E. S., Stavropoulos, T., Hoover, E. C., Eddins, D. A., Gallun, F. J., & Seitz, A. R. (2020). Portable Automated Rapid Testing (PART) for auditory assessment: Validation in a young adult normal-hearing population. *The Journal of the Acoustical Society of America, 148*(4), 1831. https://doi.org/10.1121/10.0002108

Levi, D. M., & Li, R. W. (2009). Perceptual learning as a potential treatment for amblyopia: A mini-review. *Vision Research, 49*(21), 2535–2549. https://doi.org/10.1016/j.visres.2009.02.010

Li, R. W., Klein, S. A., & Levi, D. M. (2008). Prolonged perceptual learning of positional acuity in adult amblyopia: Perceptual template retuning dynamics. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience, 28*(52), 14223–14229. https://doi.org/10.1523/JNEUROSCI.4271-08.2008

Liang, J., Zhou, Y., Fahle, M., & Liu, Z. (2015a). Limited transfer of long-term motion perceptual learning with double training. *Journal of Vision, 15*(10), 1. https://doi.org/10.1167/15.10.1

Liang, J., Zhou, Y., Fahle, M., & Liu, Z. (2015b). Specificity of motion discrimination learning even with double training and staircase. *Journal of Vision, 15*(10), 3. https://doi.org/10.1167/15.10.3

Lin, S., Mihailovic, A., West, S. K., Johnson, C. A., Friedman, D. S., Kong, X., & Ramulu, P. Y. (2018). Predicting visual disability in glaucoma with combinations of vision measures. *Translational Vision Science & Technology, 7*(2), 22. https://doi.org/10.1167/tvst.7.2.22

Lu, Z.-L., & Dosher, B. A. (2022). Current directions in visual perceptual learning. *Nature Reviews Psychology, 1*(11), 654–668. https://doi.org/10.1038/s44159-022-00107-2

Lu, Z.-L., Lin, Z., & Dosher, B. A. (2016). Translating perceptual learning from the laboratory to applications. *Trends in Cognitive Sciences, 20*(8), 561–563. https://doi.org/10.1016/j.tics.2016.05.007

Maniglia, M., Cottereau, B. R., Soler, V., & Trotter, Y. (2016a). Rehabilitation approaches in macular degeneration patients. *Frontiers in Systems Neuroscience, 10*, 107. https://doi.org/10.3389/fnsys.2016.00107

Maniglia, M., Pavan, A., Sato, G., Contemori, G., Montemurro, S., Battaglini, L., & Casco, C. (2016b). Perceptual learning leads to long lasting visual improvement in patients with central vision loss. *Restorative Neurology and Neuroscience, 34*(5), 697–720. https://doi.org/10.3233/RNN-150575

Maniglia, M., & Seitz, A. R. (2018). Towards a whole brain model of perceptual learning. *Current Opinion in Behavioral Sciences, 20*, 47–55. https://doi.org/10.1016/j.cobeha.2017.10.004

Maniglia, M., Soler, V., & Trotter, Y. (2020). Combining fixation and lateral masking training enhances perceptual learning effects in patients with macular degeneration. *Journal of Vision, 20*(10), 19. https://doi.org/10.1167/jov.20.10.19

Nuutinen, M., Mustonen, T., & Häkkinen, J. (2018). CFS MATLAB toolbox: An experiment builder for continuous flash suppression (CFS) task. *Behavior Research Methods, 50*(5), 1933–1942. https://doi.org/10.3758/s13428-017-0961-z

Pahor, A., Mester, R. E., Carrillo, A. A., Ghil, E., Reimer, J. F., Jaeggi, S. M., & Seitz, A. R. (2022). UCancellation: A new mobile measure of selective attention and concentration. *Behavior Research Methods, 54*(5), 2602–2617. https://doi.org/10.3758/s13428-021-01765-5

Pahor, A., Stavropoulos, T., Jaeggi, S. M., & Seitz, A. R. (2019). Validation of a matrix reasoning task for mobile devices. *Behavior Research Methods, 51*(5), 2256–2267. https://doi.org/10.3758/s13428-018-1152-2

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., Kastman, E., & Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods, 51*(1), 195–203. https://doi.org/10.3758/s13428-018-01193-y

Plank, T., Rosengarth, K., Schmalhofer, C., Goldhacker, M., Brandl-Rühle, S., & Greenlee, M. W. (2014). Perceptual learning in patients with macular degeneration. *Frontiers in Psychology, 5*, 1189. https://doi.org/10.3389/fpsyg.2014.01189

Polat, U. (2009). Making perceptual learning practical to improve visual functions. *Vision Research, 49*(21), 2566–2573. https://doi.org/10.1016/j.visres.2009.06.005

Polat, U., Ma-Naim, T., & Spierer, A. (2009). Treatment of children with amblyopia by perceptual learning. *Vision Research, 49*(21), 2599–2603. https://doi.org/10.1016/j.visres.2009.07.008

Polat, U., Schor, C., Tong, J.-L., Zomet, A., Lev, M., Yehezkel, O., Sterkin, A., & Levi, D. M. (2012). Training the brain to overcome the effect of aging on the human eye. *Scientific Reports, 2*, 278. https://doi.org/10.1038/srep00278

Rohr, M., & Wagner, A. (2020). How Monitor characteristics affect human perception in visual computer experiments: CRT vs. LCD monitors in millisecond precise timing research. *Scientific Reports, 10*(1), 1–10. https://doi.org/10.1038/s41598-020-63853-4

Ross, J. E., Bron, A. J., & Clarke, D. D. (1984). Contrast sensitivity and visual disability in chronic simple glaucoma. *The British Journal of Ophthalmology, 68*(11), 821–827. https://doi.org/10.1136/bjo.68.11.821

Sabour-Pickett, S., Loughman, J., Nolan, J. M., Stack, J., Pesudovs, K., Meagher, K. A., & Beatty, S. (2013). Visual performance in patients with neovascular age-related macular degeneration undergoing treatment with intravitreal ranibizumab. *Journal of Ophthalmology, 2013*, 268438. https://doi.org/10.1155/2013/268438

Sagi, D. (2011). Perceptual learning in Vision Research. *Vision Research, 51*(13), 1552–1566. https://doi.org/10.1016/j.visres.2010.10.019

Schoups, A. A., Vogels, R., & Orban, G. A. (1995). Human perceptual learning in identifying the oblique orientation: Retinotopy, orientation specificity and monocularity. *The Journal of Physiology, 483*(Pt 3), 797–810. https://doi.org/10.1113/jphysiol.1995.sp020623

Schoups, A., Vogels, R., Qian, N., & Orban, G. (2001). Practising orientation identification improves orientation coding in V1 neurons. *Nature, 412*(6846), 549–553. https://doi.org/10.1038/35087601

Seitz, A. R. (2017). *Perceptual Learning. Current Biology: CB, 27*(13), R631–R636. https://doi.org/10.1016/j.cub.2017.05.053

Sotiropoulos, G., Seitz, A. R., & Seriès, P. (2018). Performance-monitoring integrated reweighting model of perceptual learning. *Vision Research, 152*, 17–39. https://doi.org/10.1016/j.visres.2018.01.010

Talluri, B. C., Hung, S.-C., Seitz, A. R., & Seriès, P. (2015). Confidence-based integrated reweighting model of task-difficulty explains location-based specificity in perceptual learning. *Journal of Vision, 15*(10), 17. https://doi.org/10.1167/15.10.17

Wenliang, L. K., & Seitz, A. R. (2018). Deep neural networks for modeling visual perceptual learning. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience, 38*(27), 6028–6044. https://doi.org/10.1523/JNEUROSCI.1620-17.2018

Xiao, L.-Q., Zhang, J.-Y., Wang, R., Klein, S. A., Levi, D. M., & Yu, C. (2008). Complete transfer of perceptual learning across retinal locations enabled by double training. *Current Biology: CB, 18*(24), 1922–1926. https://doi.org/10.1016/j.cub.2008.10.030

Yaghoubi, K. C., Kabbara, S., Arian, S., Kobaissi, H., Peters, M. A. K., & Seitz, A. R. (2022). Comparing random dot motion in MATLAB vs inquisit millisecond. *Frontiers in Psychology, 13*, 1035518. https://doi.org/10.3389/fpsyg.2022.1035518

Yu, D., Cheung, S.-H., Legge, G. E., & Chung, S. T. L. (2010). Reading speed in the peripheral visual field of older adults: Does it benefit from perceptual learning? *Vision Research, 50*(9), 860–869. https://doi.org/10.1016/j.visres.2010.02.006

Zhang, J.-Y., & Yu, C. (2016). The transfer of motion direction learning to an opposite direction enabled by double training: A reply to Liang et al. (2015) [Review of *The transfer of motion direction learning to an opposite direction enabled by double training: A reply to Liang et al. (2015)*]. *Journal of Vision, 16*(3), 29. https://doi.org/10.1167/16.3.29

Zhang, G.-L., Cong, L.-J., Song, Y., & Yu, C. (2013). ERP P1–N1 changes associated with Vernier perceptual learning and its location specificity and transfer. *Journal of Vision, 13*(4), 19. https://doi.org/10.1167/13.4.19

Zhang, G.-L., Li, A.-S., Miao, C.-G., He, X., Zhang, M., & Zhang, Y. (2018). A consumer-grade LCD monitor for precise visual stimulation. *Behavior Research Methods, 50*(4), 1496–1502. https://doi.org/10.3758/s13428-018-1018-7

Zhang, T., Xiao, L.-Q., Klein, S. A., Levi, D. M., & Yu, C. (2010). Decoupling location specificity from perceptual learning of orientation discrimination. *Vision Research, 50*(4), 368–374. https://doi.org/10.1016/j.visres.2009.08.024