Controlling the spatial dimensions of visual stimuli in online experiments

Jan W. Brascamp

Michigan State University, Department of Psychology, East Lansing, MI, USA

There are clear benefits to using an online environment for human subjects' research, for instance, rapid data collection and access to a diverse body of potential participants. One distinct drawback of online environments as compared to laboratory environments is the relative lack of control over experiment conditions. For research into human vision, a specific concern is the relative lack of control over angular stimulus dimension in an online setting. This paper examines three approaches to estimating a participant's viewing distance online, and quantifies the magnitude of the error in angular stimulus size associated with each method. For each method, the average expected error is smaller than 20% of the intended stimulus size, and for the best method it is close to 10%. This paper provides a discussion of the benefits and drawbacks of each of the three methods, as well as parameter values and computer code that will facilitate the use of these methods in future online studies.

Introduction

Studies in the field of experimental psychology can benefit from administering experiments via online hosting services, as opposed to bringing participants into the laboratory (Reips, 2002; Crump, McDonnell, & Gureckis, 2013; Sauter, Draschkow, & Mack, 2020). This approach can increase the number participants that can feasibly be tested in a given amount of time, for instance, when combined with a university's subject pool, and especially when combined with a web-based recruitment platform, such as Prolific (Palan & Schitter, 2018) or Amazon's MTurk. The large number of potential participants that can be reached via such platforms, moreover, is a plus for studies that aim to recruit participants who meet special, low-prevalence, criteria (e.g. having a certain medical condition), or participants who form a representative sample of the overall population. In addition to such general benefits of online experiments, the restrictions in person-to-person interaction imposed by the coronavirus disease 2019 (COVID-19) pandemic have temporarily made online experiments the only option for many research laboratories around the world (Sauter et al., 2020). Indeed, it is plausible that this forced foray into online experiments will have a lasting impact on the popularity of such experiments, even in the post-pandemic world, as numerous researchers will now have climbed the learning curve of online testing.

A potential limitation of trading in the laboratory computer for participants' own personal machines, is a relative lack of control over experiment conditions. In the present study, I am concerned with control over the participant's viewing distance and, ultimately, over the angular dimensions of on-screen stimuli. Many classic phenomena in cognitive psychology are robust to the relative lack of control associated with online testing (Germine, Nakayama, Duchaine, Chabris, & Chatterjee, 2012; Crump et al., 2013), but uncertainty about viewing distance can be a serious limitation, particularly in the field of vision science, where the angular extent and angular eccentricity of stimuli on the retina are often critical. Indeed, in vision science, these variables are routinely specified in methods sections and are typically kept constant in the laboratory with the use of a head rest. There is good reason for this, because countless (probably most) aspects of visual function show a systematic dependence on angular stimulus size and eccentricity (e.g. Rubin, Nakayama, & Shapley, 1997; Pelli, Palomares, & Majaj, 2004; Tadin & Lappin, 2005; Kang, 2009; Herrmann, Montaser-Kouhsari, Carrasco, & Heeger, 2010; Stewart, Valsecchi, & Schütz, 2020). Clearly, then, the value of online experiment environments for vision-related research depends critically on the experimenters' ability to infer or control viewing distance and, consequently, angular stimulus dimensions.

In this study, I conducted two online experiments that were administered via a web-based recruitment platform, in order to compare three approaches to estimating viewing distance online. First, I evaluated how accurately one can estimate viewing distance from a participant's response to two simple questions: whether the participant is using a laptop screen or an external screen, and how tall the participant is. This first approach is based on the common-sense assumptions that the range of natural viewing distances is limited to

Citation: Brascamp, J. W. (2021). Controlling the spatial dimensions of visual stimuli in online experiments. *Journal of Vision*, *21*(8):19, 1–11, https://doi.org/10.1167/jov.21.8.19.

https://doi.org/10.1167/jov.21.8.19 Received May 20, 2021; published August 19, 2021 ISSN 1534-7362

begin with, and that it depends on what type of screen a person uses (closer for a laptop screen) as well as on the person's height (closer for shorter people, particularly when using a laptop). Second, I evaluated an approach that asks participants to position themselves at arm's length from the screen center. In combination with the participant's height, this may provide a reasonable estimate of viewing distance, given that body height and arm length are closely correlated (Quanjer, Capderou, Mazicioglu, Aggarwal, Banik, Popovic, Tayie, Golshan, Ip, & Zelter, 2014). Third, I compared those two methods to a recently proposed method that centers on estimating the position of the blind spot on the screen (Li, Joo, Yeatman, & Reinecke, 2020). The angular eccentricity of the blind spot is relatively constant across participants, which allows its screen position to be translated into a fairly reliable measure of viewing distance – an approach that Li et al. (2020) termed the "virtual chinrest" method.

To preview the results, I find that the method based on blind spot location allows the most accurate control of angular stimulus dimensions, followed by the method based on arm length. I also find that the method based only on two questions, although not as good as the other two methods, does not perform dramatically worse and may be suitable in some situations where only approximate control of angular dimensions is required or where an alternative approach fails. I provide parameter values that can be used to translate observed variables (body height, screen type, and blind spot location) into estimates of viewing distance in future experiments, and I report the magnitude of the error in estimated distance and in angular stimulus dimensions that can be expected when using these parameter values. Finally, based on the observation that participant responses are not always reliable, I propose a tiered approach that attempts several approaches to estimate viewing distance in sequence. I provide code created in PsychoPy (Peirce, Gray, Simpson, MacAskill, Höchenberger, Sogo, Kastman, & Lindeløv, 2019) to implement this tiered approach at the start of an online experiment.

Methods

I performed two experiments. Both were created in PsychoPy (Peirce et al., 2019) and hosted on its online platform Pavlovia. Participants were recruited and paid using Prolific (Palan & Schitter, 2018). All experiments were approved by the Institutional Review Board of Michigan State University. In the first experiment, I tested the approach based on body height and screen type, and the approach based on arm length. In the second experiment, I tested the approach based on blind spot location.

Experiment 1

For this experiment, I recruited 33 colleagues from vision science and related fields, as well as 421 participants from the general population. The 33 colleagues were recruited via email and volunteered their time. The remaining participants were recruited using Prolific, and received $1.59 for a 10-minute experiment ($9.54 per hour). Before the experiment started participants were instructed to get a tape measure or equivalent, which they would need during the experiment. The experiment started with a few questions, asking the participant's race, age, sex assigned at birth, body height, and screen type (laptop screen or external screen). I considered that the factors of race, age, and sex could be relevant because they might moderate, for example, the relation between body height and arm length, but preliminary analyses did not support this idea (data not shown) so those factors will not be discussed further. After the initial questions, the experiment asked to "position yourself naturally in front of this computer, just as you normally would when using this computer." Then, while instructing the participant to remain in that position, the experiment asked the participant to measure the distance between one eye and a marker at the screen center. Finally, participants were asked to position themselves at arm's length from the screen (Figure 1) and then to measure that distance again. For the 421 participants from the

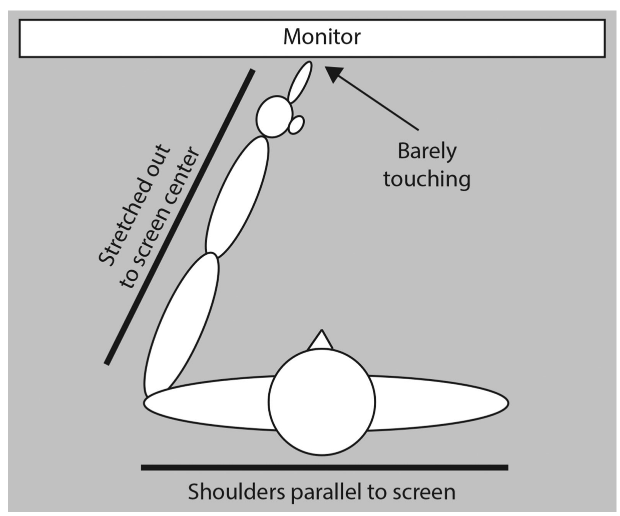

Figure 1. The image used to help participants position themselves at arm's length from the screen. The accompanying instructions read as follows: "Reach forward with one of your arms, but don't rotate your shoulder forward. Instead, your left shoulder should be just as far from the screen as your right shoulder. Then sit at such a distance from your screen that you can just barely touch the red dot at the center of your screen with your index finger, while keeping your arm, hand, and finger stretched out straight to the red dot."

general population (but not for the 33 colleagues), there was one additional task, right after the initial questions and before the first distance measure. This task involved measuring the length of a line on the screen, and the purpose of the task was to make sure that the participant had, indeed, found a tape measure and knew how to use it. Because this required the experimenter knowing the actual length of the line, the participants from the general population also performed a procedure, at the very start of the experiment, that involves scaling an on-screen picture to the size of a bank card (see figure 1a in Li et al., 2020, and https://gitlab.pavlovia.org/Wake/screenscale), which allows the experimenter to infer the size of a pixel on each participant's screen. I did not find any significant differences between the colleagues and the participants from the general population, so those participant groups are pooled together in all analyses presented here.

I excluded all men who reported a body height smaller than 153.5 cm or larger than 203.5 cm, and for women those limits were 141.5 cm and 188 cm, respectively (based on the 0.05% and 99.95% points of the world body height distributions for 20-year-olds found on https://www.gigacalculator.com/calculators/height-percentile-calculator.php). I also excluded all participants who entered the length, in whole units, of the on-screen line as anything else than 18 cm or 7 inches (depending on which unit of preference they selected during the experiment). The physical length of the line was 18 cm, which translates to 7.09 inches. These criteria combined led to the exclusion of three colleagues (9%) and of 196 of the remaining participants (47%; 32 failed the height criterion, 130 failed the line length criterion, and 34 failed both). Of the remaining participants, I excluded a final eight from further analysis because they indicated a natural viewing distance and/or arm's length viewing distance that lay more than three standard deviations from the relevant average. (For arm's length viewing distance, these average and standard deviations were computed across all participants not excluded based on earlier criteria; for natural viewing distance they were computed for each screen type separately.) This left 247 participants for the main analyses.

Experiment 2

For this experiment, I used Prolific to recruit 403 participants from the general population. They received \$1.59 for a 10-minute experiment (\$9.54 per hour). From the experiment components listed above, this experiment involved the credit card scaling procedure, the questions about body height and screen type (but not the other questions), and the tasks of measuring on-screen line length and screen distance at one's natural position (but not the task of measuring the distance at arm's length). After that, the participant completed a task, not included in experiment 1, meant to find the location of the blind spot. Similar to tasks used elsewhere for this purpose (e.g. Li et al., 2020), this task involved participants closing or covering their right eye and directing their gaze at a central fixation point, while a blue square (side 0.5 cm) repeatedly moved away from and back toward fixation, sliding horizontally at fixation level on the left side of the screen. The square started at its most central position, at an intended eccentricity of 8 degrees of visual angle (dva) and moved back and forth between that position and its most peripheral position, which was at an intended eccentricity of 17 dva, or at 0.012 screen widths inside the edge of the screen, whichever was smaller. As soon as the square reached its most extreme position, it reversed direction, moving back and forth a total of three times (i.e. it moved in each direction 3 times), at a constant speed such that one passage took 7 seconds. For this procedure, the intended angular eccentricities were translated to on-screen coordinates based on the viewing distance as measured and reported by the participant just earlier. Based on the range of viewing distances observed in experiment 1 (see below), I only accepted reported viewing distances between 30 cm and 110 cm for this particular purpose, and interpreted values outside of that range as erroneous and replaced them with the average of the range: 70 cm. Participants were instructed to hold down the spacebar whenever they saw the blue dot, and to release it whenever they did not. The objective was to find the position, in terms of centimeters on the screen, of the inner or nasal edge of the blind spot, which an earlier publication placed at 13.6 dva on average, with a standard deviation of 0.96 dva (Li et al., 2020). Although the actual angular eccentricity of the blue square during this procedure depended on the reliability of the participant's earlier report of viewing distance, the procedure was fairly robust to errors in that report, given the broad range of intended angular positions (from 8 to 17 dva) relative to that published distribution of blind spot edge positions (13.6 ± 0.96 dva).

In an offline analysis of the blind spot data, I divided each participant's data into six passages of the blue square: three toward the periphery and three toward the fovea (see Figure 2). For each passage, I identified periods during which no spacebar was registered for at least 750 ms, thus ignoring shorter periods that may separate keyboard signals while a key is being held down (see caption of Figure 2 for details). For outward-moving passages, I then marked the square's position at the time of the last recorded spacebar signal prior to the key release period (see Figure 2A). For inward moving passages, I marked the first recorded spacebar signal following the key release period (Figure 2B). Passages during which no

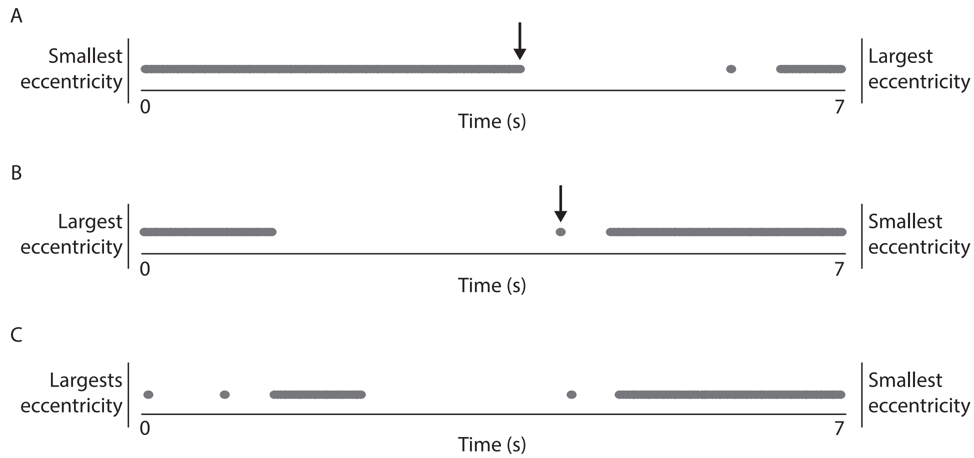

Figure 2. Example keypress data from the blind spot procedure, illustrating the first steps of the offline blind spot analysis. The target moved back and forth repeatedly during the procedure, and the resulting key press data were separated into individual periods of uninterrupted outward movement or inward movement. Each panel here shows such a period. The gray marks indicate moments at which a spacebar was registered. The downward-pointing arrows indicate moments marked as ones at which the target entered the blind spot during an outward passage (**A**), or at which it left the blind spot during an inward passage (**B**). No moment was marked for passages during which the spacebar was released multiple times (**C**). It was found that, while the spacebar was continuously held down, the interval between registered key signals was usually about 17 ms, corresponding to 60 Hz. However, this interval was often longer immediately after the first key signal had been registered upon a new key press, as is visible in all panels shown here (e.g. at the arrow in panel **B**). For this reason, the threshold for counting a period without key signals as a key release period was set well above 17 ms.

key was pressed and passages with more than one key release period were not included in further calculations (Figure 2C). All positions marked in this automated fashion were verified by visual inspection of the data, which led to manual adjustments based on subjective judgment (i.e. removal of a marked time point or addition of a missed one) in about 20 cases (i.e. for about 5% of the observers). I then discarded the data of participants for whom a position had been marked for fewer than four passages. For a given participant, I then averaged all positions marked in this fashion across all outward-moving passages, and separately across all inward-moving passages. The estimate of the blind spot's inner boundary for that participant, finally, was the average of those two averages; an approach that aims to cancel out any delays related to manual response time, which have opposite effects on the position estimate for outward-moving passages versus inward-moving ones.

I discarded data from a total of 266 participants (66%) based on either the above-mentioned blind spot criterion (78 participants), incorrectly reported line length (56 participants; same criterion as in experiment 1), a reported body height outside of the accepted range (21 participants; this range was now set from 141.5 cm to 203.5 cm for all participants because I did not record sex), or a combination of multiple of those criteria (111 participants). One final participant was removed because he/she indicated a natural viewing distance that lay more than three standard deviations from the average for the relevant screen type. This leaves 136 participants whose results are discussed below.

Results

In experiment 1, 144 participants indicated using a laptop screen and 103 participants indicated using an external screen. Figure 3A shows natural viewing distance as reported during that experiment, separately for the different screen types. Natural viewing distance clustered fairly narrowly near 60 cm (the boxes delineate the center 50% of the data and the line inside each box indicates the median; see Figure 3A caption for details), and it differed significantly between screen types (means = 51.7 cm and 61.5 cm, $t(245) = 6.54$, $p < 0.0001$). Arm's length viewing distance (see Figure 3A) appeared a little larger and more narrowly clustered, and was not significantly different between screen types (means = 64.1 cm and 62.8 cm, $t(245) = 0.90$, $p = 0.37$).

Body height significantly predicted natural viewing distance for laptop users but not for participants with an external screen (Figure 4A), and it predicted arm's length viewing distance regardless of screen type (Figure 4B). Comparing those results side-by-side with the results of experiment 2 (Figure 4C) shows that the relation between natural viewing distance and on-screen blind spot location is substantially stronger than either of those two relations, with strong

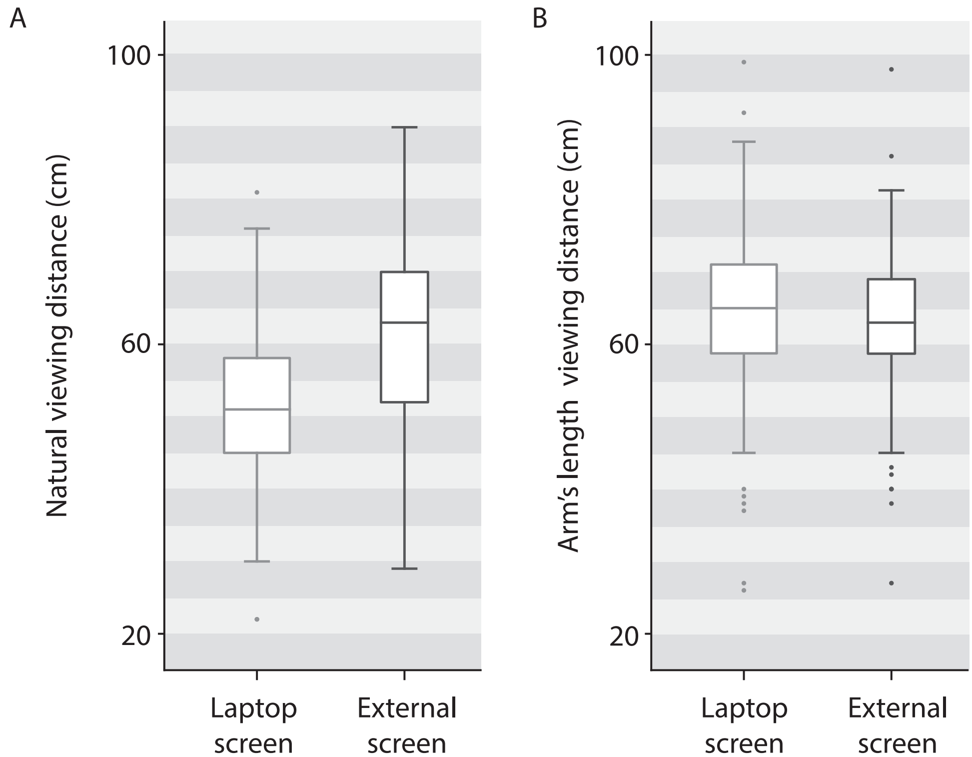

Figure 3. Viewing distance per screen type, when positioned naturally (**A**) or when positioned at arm's length (**B**). Each box contains the central 50% of the data (ranging from the first to the third quantile) with the median marked by a horizontal line. Box width is proportional to the number of data points in the corresponding data set (144 and 103 for laptop screens and for external screens, respectively). The two whiskers attached to a box extend, respectively, to the most extreme sub-median data point and the most extreme supra-median data point within 1.5 box heights from the box. The remaining data points are plotted individually and would conventionally be considered outliers.

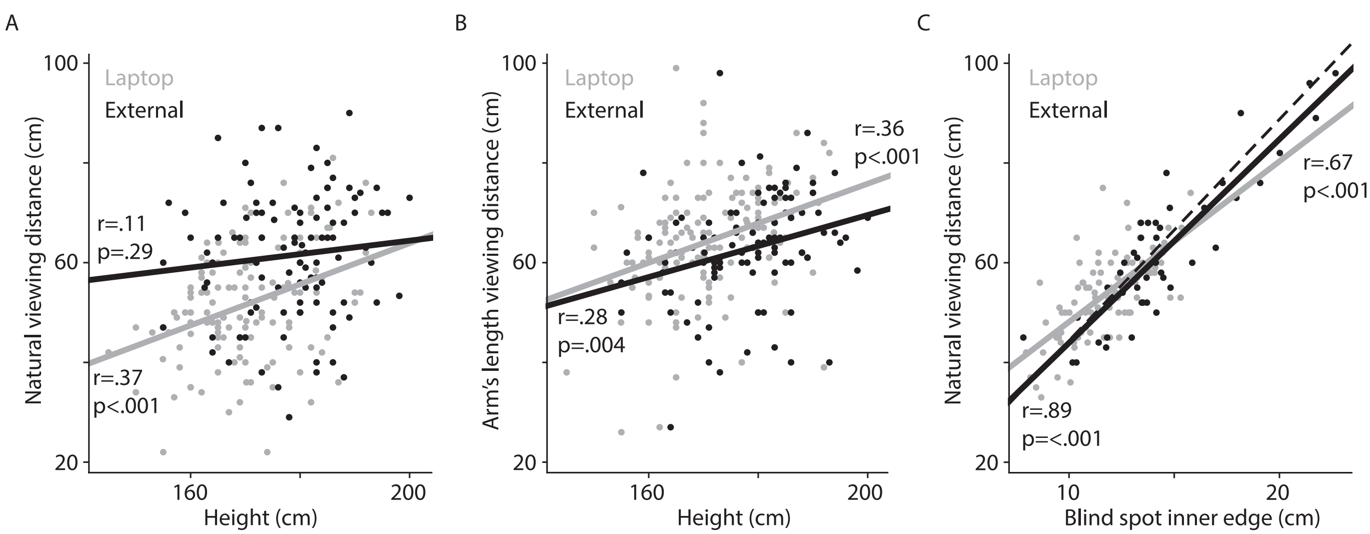

Figure 4. Relation between predictor variables (body height and on-screen blind spot position) and viewing distance, separated out by screen type. Each solid line represents a two-parameter linear fit. (**A**) Natural viewing distance as a function of body height. (**B**) Arm's length viewing distance as a function of body height. (**C**) Natural viewing distance as a function of on-screen blind spot position. The dashed line in panel **C** corresponds to a one-parameter linear fit (with the offset fixed at 0) to data across both screen types.

	Offset (cm)	Slope (no unit)
Natural distance vs. height, laptop	−18.5	0.41
Natural distance vs. height, external	37.4	0.13
Natural distance vs. height, screen types combined	−18.3	0.43
Arm's length distance vs. height, laptop	−3.5	0.40
Arm's length distance vs. height, external	7.4	0.31
Arm's length distance vs. height, screen types combined	13.4	0.29
Natural distance vs. blind spot position, laptop	15.9	3.22
Natural distance vs. blind spot position, external	3.1	4.08
Natural distance vs. blind spot position, screen types combined	11.4	3.57
Natural distance vs. blind spot position, screen types combined, no offset	N/A	4.44

Table 1. Fit parameters relating observed variables to viewing distance.

and significant correlations both for laptop users (91 participants) and for participants with an external screen (45 participants). Figure 4C also shows a third curve in addition to the two solid curves that were also shown in the other panels. Whereas the solid curves in these panels correspond to linear fits with two free parameters, both offset and slope, this dashed curve corresponds to a fit for which the offset was pegged at zero. If, as has been reported in earlier studies, the angular eccentricity of the (inner edge of the) blind spot is approximately constant across observers (Rohrschneider, 2004; Wang, Shen, Boland, Wellik, De Moraes, Myers, Bex, & Elze, 2017; Li et al., 2020), then the slope of this dashed line can be converted into this angular eccentricity. Indeed, the inner edge eccentricity that corresponds with the slope found here is 12.7 dva, in good concordance with the estimates provided in those studies.

Table 1 shows all parameters of the fits depicted in Figure 4 and also of additional related fits. The values in this table will allow researchers to estimate viewing distance based on a participant's screen type, body height, and/or on-screen blind spot position in future experiments.

The results shown so far give an indication of the viewing distance of participants in online experiments and how it relates to variables that can feasibly be collected in such experiments. The key question, however, is: when basing oneself on those variables, how large an error may one expect in the estimated viewing distance and, more importantly, in the angular dimensions of a stimulus on the screen? Figure 5 is concerned with that question.

Figure 5A shows the error in estimated viewing distance for each of the three approaches (natural viewing distance based on height, arm's-length viewing distance based on height, and natural viewing distance based on blind spot location). In each case, the distance estimate for a given participant was based on a two-parameter linear fit to the data from all participants with the same screen type as the participant in question, except the participant himself/herself (to avoid circularity). The error was then calculated by comparing the distance estimate read off of that fitted curve to the viewing distance actually reported by the participant. Each box in Figure 5A contains the center 50% of all data points, with the line inside the box indicating the median. The box heights show that, for each method, the error in estimated distance is smaller than 10 cm for at least half of the participants. The whiskers delimit all data that lie within a certain number of box heights from a given box (see caption for details). In this case, this means that the whiskers delimit between 94% and 99% of the data, depending on the method. All remaining data points are plotted individually. Comparing between methods, Figure 5A suggests relatively large errors for the method that involves natural distance and body height (left), and particularly small errors for the method that involves the blind spot (right). For this latter method, the best 50% of distance estimates are all within 5 cm of the reported distance, and nearly all participants have a distance estimate error of well under 20 cm. This finding of good performance for the latter method is in full agreement with the report by Li et al. (2020).

The same relative ordering in the performance of the three methods is observed when analyzing the absolute, rather than signed, values of the distance estimate errors. For the three methods, the average absolute errors are 8.9 cm ("height and natural"), 7.6 cm ("height and arm's length"), and 5.5 cm ("blind spot and natural"), and the difference between methods is significant (ANOVA: $F(2,628) = 11.53$, $p < 0.0001$; paired t-test between the two height-based methods: $t(246) = 2.24$, $p = 0.03$; t-test between "height and natural," and "blind spot and natural": $t(380) = 5.18$, $p < 0.0001$; t-test between "height and arm's length" and "blind spot and natural": $t(380) = 2.99$, $p < 0.01$).

More important than inaccuracies in estimated viewing distance are inaccuracies in the angular dimensions of on-screen stimuli that are drawn on the basis of those distance estimates. Figure 5B

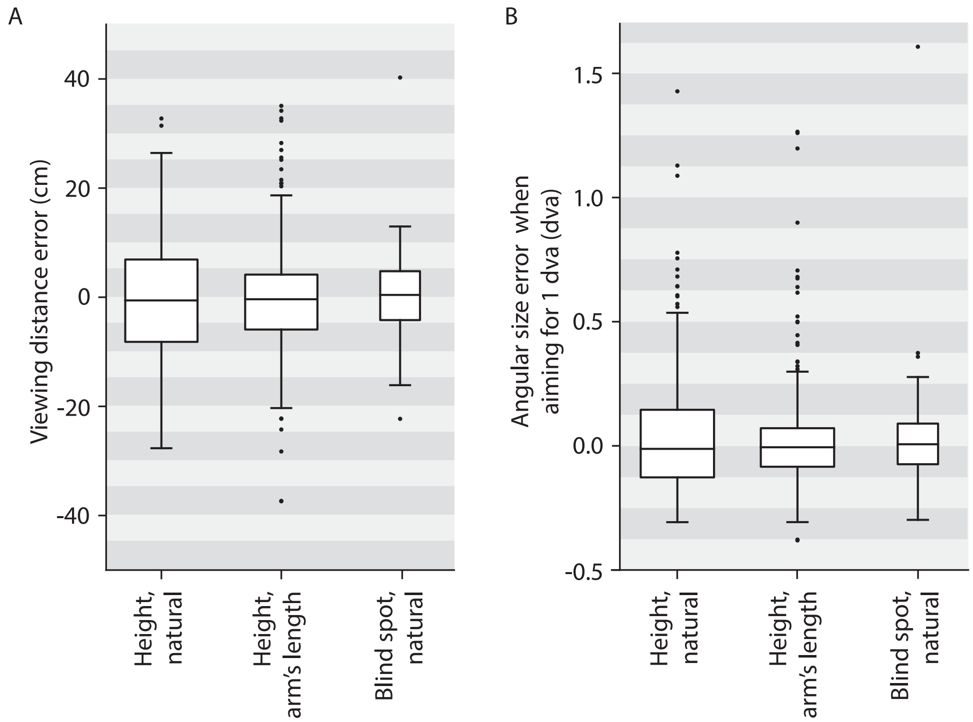

Figure 5. Error in estimated distance (**A**) and in the angular extent of a stimulus that is drawn on the basis of the estimated distance (**B**). Each box contains the central 50% of the data (ranging from the first to the third quantile) with the median marked by a horizontal line. Box width is proportional to the number of data points in the corresponding data set (247 for the left and center box in each panel; 136 for the right box). The two whiskers attached to a box extend, respectively, to the most extreme sub-median data point and the most extreme supra-median data point within 1.5 box heights from the box. The remaining data points are plotted individually and would conventionally be considered outliers. Going from left to right, these outliers comprise, respectively, 0.8%, 6.4%, and 1.5% of the data for panel **A**, and 4.9%, 8.5%, and 2.2% of the data for panel **B**. Note that panel **B**, although labeled as showing angular size error in dva, can equivalently be read as showing size error as a proportion of intended size (because intended size here is 1 dva). Also note that computing viewing distance error as a proportion of actual distance (not shown here) would yield values almost identical to those of panel **B**.

summarizes, for each method, the amount by which the radius of a stimulus at the screen center would differ from its intended value, if the experimenter drew the stimulus based on the estimated viewing distance, and if the intended radius was 1 dva. The y-axis of this panel is labeled as quantifying this error in terms of degrees of visual angle, so it can equivalently be read as quantifying the error as a proportion of the intended radius. When expressed as a proportion, the error is largely independent of intended stimulus size within the size range that is relevant to most vision experiments (within that range proportional error becomes very slightly smaller as a function of intended stimulus size). The box plots are again constructed in the same way as before, but this time that works out to between 91% and 98% of the data being contained between the whiskers (see caption for details). Figure 5B shows that the proportional error in angular extent is comparable between the "height and arm's length" method (center) and the blind spot method (right). In both cases, the proportional error in angular extent lies well under one-eighth for at least half of the participants (i.e. the boxes lie within one gray band on either side of 0), and rarely exceeds three-eighths (i.e. the whiskers lie within 3 gray bands on either side of 0). For the "height and natural" method, the errors appear larger: the box extends slightly beyond the one-eighth mark in the positive direction, and the whisker in that direction reaches just beyond four-eigths, which indicates that it is relatively common for stimuli to be drawn as much as 50% larger than they should be when using this method. A final observation is that there are many more outliers (individually plotted data points) in the positive direction than in the negative direction in this panel, which means that the distribution is skewed such that is more common to draw the on-screen images much too large, in terms of angular extent, than it is to draw them much too small. This is related to the geometric fact that, for a given distance estimate, the error in angular stimulus extent will be relatively large (and positive) if that estimate overshoots the real distance by a given amount, but relatively small (and negative) if the estimate undershoots the real distance by the same amount.

An analysis of the absolute value of the errors again confirms the relative ordering in performance of the three methods that is suggested by the figure. For the three methods, the average absolute errors in stimulus extent are 0.18 dva ("height and natural"), 0.14 dva ("height and arm's length"), and 0.11 dva ("blind spot and natural"). Equivalently, these average absolute errors are 0.18 times the intended stimulus size, 0.14 times the intended stimulus size, and 0.11 times the intended stimulus size, respectively. The difference between methods is overall significant (ANOVA: $F(2,628) = 7.65$, $p < 0.001$; paired t-test between the two height-based methods: $t(246) = 2.99$, $p = 0.003$; t-test between "height and natural" and "blind spot and natural": $t(380) = 3.80$, $p < 0.001$), but the difference between the arm's length method and the blind spot method is not significant ($t(380) = 1.55$, $p = 0.12$).

Discussion

I examined three approaches to estimating viewing distance and, related to that, to controlling angular stimulus dimensions in online experiments. Consistent with a previous report (Li et al., 2020) I found that viewing distance can be estimated with a fair degree of accuracy by measuring the on-screen position of the blind spot. I also found that asking participants to remain at arm's length from the screen, and using their body height to estimate viewing distance, allows angular stimulus extent to be controlled with a level of accuracy that is very similar to what is allowed by the blind spot method. Finally, I found that viewing distance can be estimated, albeit with a lower degree of accuracy, based on nothing but knowledge about the participant's body height and screen type (laptop or external). I provide parameter values that allow future researchers to translate observed variables into viewing distance estimates in their studies, and I quantify the magnitude of the error in angular stimulus dimensions that may be expected when using that approach.

Aside from the quantitative comparison between methods that is presented in this paper, there are further considerations that affect which approach a researcher may take. Compared to the other two methods, a drawback of the method that asks participants to position themselves at arm's length from the screen, is that this may not be the participant's natural position. This plausibly increases the likelihood for viewing distance to change during the experiment. In contrast, the other two methods allow the participant to stay in their natural position, at a viewing distance of their own choice, which plausibly results in a reasonably constant viewing distance over time, even without the possibility of using a head rest such as would be used in a laboratory environment. An estimate of the extent of viewing distance variation in conditions like these comes from the study by Li et al. (2020). One of their online experiments, in which participants were asked to stay in a constant, comfortable position, included three separate repetitions of the blind spot estimation procedure, spaced apart by about 5 minutes during which 25 trials of a visual crowding task were administered. The authors found that estimated viewing distance had an average within-participant standard deviation of 3.9 cm across the three repetitions. In general, it could be advisable to estimate viewing distance several times during an online experiment, both to adjust display dimensions to small changes in viewing distance, and to be able to discard data from participants who do not sit still.

A further benefit of the method based on the blind spot is that it does not necessarily require the researcher to know the dimensions of a pixel on the participant's screen. As detailed above, these dimensions can be estimated via a procedure in which the participant compares on-screen dimensions to those of a real-world object that has a standard size, such as a bank card. However, this procedure can be omitted when estimating viewing distance from a blind spot position on the basis of only a slope parameter, such as the one documented in the bottom row of Table 1 (i.e. the slope of the fit with an offset of zero, corresponding to a fixed angular blind spot eccentricity for all participants). Specifically, in this approach, the researcher can simply relate blind spot eccentricity in pixels to viewing distance, also measured in pixels, using the slope parameter. That distance, in turn, can then be used to translate angular extent to on-screen dimensions, again in pixels, without knowing the centimeter value associated with any of these quantities. Given that the approach that centers on the blind spot uses only its horizontal eccentricity, in the absence of knowledge about pixel dimensions, the approach strictly allows control over angular stimulus extent only in the horizontal direction. In practice, however, the vast majority of computer screens have square pixels (at least the screens used by participants in the present study: out of 353 included participants who completed the credit card procedure, 345, or 97.7%, indicated a vertical/horizontal pixel aspect ratio between 0.95 and 1.05). Pixel aspect ratio is also something that can be verified in a matter of seconds by, say, asking participants to choose which of several ellipses drawn on the screen is round.

Several findings in the present study constitute a replication of findings reported by Li et al. (2020). That study focused on the relation between on-screen blind spot position and viewing distance, and it did not involve the other two methods examined here. The study also did not include any experiments that involved online participants measuring their own viewing distance (the study did include an online component, but relied on in-laboratory experiments for analyses that required a ground-truth estimate of distance). To the extent that the present findings overlap with those of that prior study, they are in good agreement. One slight difference is in the estimate of the angular eccentricity of the blind spot's nasal boundary at the horizontal meridian, which I put at 12.7 dva, compared to 13.6 dva in Li et al. (2020). One possible explanation for this difference is that Li and collaborators' estimate is based on only outward movement of the visual target. Given the manual reaction time delay involved in reporting a target's apparent disappearance (or reappearance), an approach based only on outward movement likely results in some degree of overestimation of the blind spot's eccentricity. The present approach involves averaging an estimate obtained during outward movement and a separate estimate obtained during inward movement, thus sidestepping that issue. Consistent with this explanation, Li et al. (2020) report their estimate of the blind spot's eccentricity (15.8 dva when expressed in terms of the blind spot center rather than its nasal edge) to be about one degree larger than the estimate provided by other studies that used different methods (14.3–15.5 when expressed in that way), which means that the present estimate matches the estimates of those other studies well.

A class of approaches to viewing distance estimation that is not examined in the present study, is the class that centers on face tracking using a webcam. Webcams are ubiquitous in modern-day personal computers as well as in mobile devices, and software for delineating faces and their features inside a webcam image is available to the general public (Kartynnik, Ablavatski, Grishchenko, & Grundmann, 2019; https://github.com/tensorflow/tfjs-models/tree/master/facemesh; https://github.com/auduno/clmtrackr). Webcam-based facial feature detection can conceivably play either an auxiliary role or a central role in viewing distance estimation. In particular, if the experimenter has no knowledge of the absolute sizes of the facial features, then the video stream could still be used to identify relative changes in viewing distance, which could inform updates to a previously obtained viewing distance estimate, or trigger a new estimation procedure. If the experimenter does have knowledge of the absolute sizes of the facial features, then the video stream can, itself, provide the basis for a real-time distance estimate. Such absolute knowledge could come from a procedure akin to the credit card procedure used here (see Methods) but where the participant holds an object of known size by his/her face in the camera image. Alternatively, it could come from anatomic constants. Indeed, an application exists that uses the fact that iris diameter is relatively preserved across participants, to estimate viewing distance from camera data with considerable accuracy (https://ai.googleblog.com/2020/08/mediapipe-iris-real-time-iris-tracking.html). When implemented in online experiment software, and with the provision that they require consent to access the participant's camera, such video-based methods can form a highly effective approach to viewing distance estimation.

One final, but important, consideration when it comes to the utility of the approaches studied here, is attrition. A sizable proportion of participants failed this study's inclusion criteria. Many of those exclusions had to do with the participants' own manual distance measurements, and are therefore irrelevant to the utility of the approaches under study, as those do not involve such measurements. Still, a substantial number of exclusions remains when considering only exclusion criteria that are inherent to the approaches under study. In experiment 1, about 15% of the participants reported an implausible body height. In experiment 2, that percentage was 5%, whereas for a further 20% of participants no blind spot location could be inferred, and a final 6% failed both those inclusion criteria. There are several ways of reducing the number of excluded participants. For instance, the experiment can be programmed to accept only plausible body height values, and in the blind spot procedure the visual target can be programmed to keep moving back and forth until a satisfactory estimate of blind spot location is obtained. On the other hand, it is plausible that some proportion of the participants who fail these inclusion criteria are generally unmotivated or otherwise unlikely to produce high quality data (indeed, in the present study, the proportion of participants who failed multiple criteria is much higher than what would be expected based on each criterion's individual proportion when assuming independence). From that perspective, it makes sense to allow participants to enter, say, an implausible body height, and to treat the question as a type of "filter" or "screening" question; the inclusion of such questions in online studies has been proposed elsewhere (e.g. Reips, 2002; Chandler, Rosenzweig, Moss, Robinson, & Litman, 2019). One final question here concerns the extent to which the error in on-screen stimulus size would increase if a researcher were to omit the stringent exclusion criteria that we applied. Clearly, obtaining a blind spot-based distance estimate is simply not an option for those observers whose blind spot location could not be inferred, but one could include height-based distance estimates even for observers who report an implausible body height. If we take that approach in our analysis of experiment 1, then the average absolute error in the radius of an on-screen object with an intended radius of 1 dva increases from 0.18 dva to 0.25 dva for the "height and natural" method. For the "height and arm's length" method, in turn, that error increases from 0.14 dva to 0.18 dva. These changes are not negligible, so keeping a close eye on data quality is recommended.

In order to benefit from the relative accuracy of the blind spot-based method (when it is successful), while simultaneously reducing the concern of data loss associated with the relative difficulty of obtaining a blind spot estimate, I propose a tiered approach. A Psychopy/Pavlovia implementation of this approach can be found at https://gitlab.pavlovia.org/janbrascamp/blind_spot, to be used and adapted as part of online experiments. The basic logic of the approach is as follows. The participant reports body height and screen type, and performs a screen scaling task (using a bank card) as well as a task to determine blind spot location. An inferred viewing distance is then computed, to be used to control stimulus dimensions in subsequent experiment code. This inference is based on estimated blind spot position if a credible estimate is obtained. If not, then viewing distance is inferred from body height and screen type, as long as the entered body height is plausible. If not, then viewing distance is estimated from assumed body height and reported screen type. The code stores the results of individual steps in the process (e.g. reported body height, estimated blind spot position for each passage of the visual target, etc.) so that, in post hoc analyses, one can decide for each participant whether the viewing distance estimate used has been satisfactory.

Conclusion

Although online environments have clear benefits as a tool for experimental psychology, they also have drawbacks, including a relative lack of control over experiment conditions. Within the context of vision research, one particular concern may be the relative lack of control over angular stimulus dimensions. This paper reports on three approaches to estimating viewing distance and, thereby, controlling stimulus dimensions in an online environment, and it provides parameter values and experiment code that will help use these approaches in future research.

Keywords: web-based experiments, online experiments, blind spot, stimulus size, methods

Acknowledgments

Commercial relationships: none.
Corresponding author: Jan W. Brascamp.
Email: brascamp@msu.edu.
Address: Michigan State University, Department of Psychology, 317 Physics Road, East Lansing, MI 48824, USA.

References

Chandler, J., Rosenzweig, C., Moss, A. J., Robinson, J., & Litman, L. (2019). Online panels in social science research: Expanding sampling methods beyond Mechanical Turk. *Behavior Research Methods, 51*, 2022–2038.

Crump, M. J. C., McDonnell, J. V., & Gureckis, T. M. (2013). Evaluating Amazon's Mechanical Turk as a Tool for Experimental Behavioral Research. *PLoS One, 8*, e57410.

Germine, L., Nakayama, K., Duchaine, B. C., Chabris, C. F., & Chatterjee, G. (2012). Is the Web as good as the lab? Comparable performance from Web and lab in cognitive/perceptual experiments. *Psychonomic Bulletin and Review, 19*, 847–857.

Herrmann, K., Montaser-Kouhsari, L., Carrasco, M., & Heeger, D. J. (2010). When size matters: attention affects performance by contrast or response gain. *Nature Neuroscience, 13*, 1554–1559.

Kang, M. S. (2009). Size matters: A study of binocular rivalry dynamics. *Journal of Vision, 9*, 17.

Kartynnik, Y., Ablavatski, A., Grishchenko, I., & Grundmann, M. (2019). Real-time Facial Surface Geometry from Monocular Video on Mobile GPUs [published online ahead of print July 15, 2019], https://arxiv.org/abs/1907.06724.

Li, Q., Joo, S. J., Yeatman, J. D., & Reinecke, K. (2020). Controlling for Participants' Viewing Distance in Large-Scale, Psychophysical Online Experiments Using a Virtual Chinrest. *Scientific Reports-UK, 10*, 904.

Palan, S., & Schitter, C. (2018). Prolific.ac—A subject pool for online experiments. *Journal of Behavioral and Experimental Finance, 17*, 22–27.

Peirce, J. W., Gray, J. R., Simpson, S., MacAskill, M. R., Höchenberger, R., Sogo, H., Kastman, E., . . . Lindeløv, J. (2019). PsychoPy2: experiments in behavior made easy. *Behavior Research Methods, 51*, 195–203.

Pelli, D. G., Palomares, M., & Majaj, N. J. (2004). Crowding is unlike ordinary masking: Distinguishing feature integration from detection. *Journal of Vision, 4*, 12.

Quanjer, P. H., Capderou, A., Mazicioglu, M. M., Aggarwal, A. N., Banik, S. D., Popovic, S., Tayie, F. A. K., Golshan, M., Ip, M. S. M., . . . Zelter, M. (2014). All-age relationship between arm span and height in different ethnic groups. *European Respiratory Journal, 44*, 905–912.

Reips, U.-D. (2002). Standards for Internet-Based Experimenting. *Exp Psychology Former Zeitschrift Für Exp Psychologie, 49*, 243–256.

Rohrschneider, K. (2004). Determination of the Location of the Fovea on the Fundus. *Investigative Ophthalmology & Visual Science, 45*, 3257–3258.

Rubin, N., Nakayama, K., & Shapley, R. (1997). Abrupt learning and retinal size specificity in illusory-contour perception. *Current Biology: CB, 7*, 461–467.

Sauter, M., Draschkow, D., & Mack, W. (2020). Building, Hosting and Recruiting: A Brief Introduction to Running Behavioral Experiments Online. *Brain Sciences, 10*, 251.

Stewart, E. E. M., Valsecchi, M., & Schütz, A. C. (2020). A review of interactions between peripheral and foveal vision. *Journal of Vision, 20*, 2–2.

Tadin, D., & Lappin, J. S. (2005). Optimal size for perceiving motion decreases with contrast. *Vision Research, 45*, 2059–2064.

Wang, M., Shen, L. Q., Boland, M. V., Wellik, S. R., De Moraes, C. G., Myers, J. S., Bex, P. J., . . . Elze, T. (2017). Impact of Natural Blind Spot Location on Perimetry. *Scientific Reports-UK, 7*, 6143.