Evaluating comparative and equality judgments in contrast perception: Attention alters appearance

Katharina Anton-Erxleben,
Psychology and Neural Science, New York University, New York, NY, USA

Jared Abrams, and
Psychology and Neural Science, New York University, New York, NY, USA

Marisa Carrasco
Psychology and Neural Science, New York University, New York, NY, USA

Abstract

Covert attention not only improves performance in many visual tasks but also modulates the appearance of several visual features. Studies on attention and appearance have assessed subjective appearance using a task contingent upon a comparative judgment (e.g., M. Carrasco, S. Ling, & S. Read, 2004). Recently, K. A. Schneider and M. Komlos (2008) questioned the validity of those results because they did not find a significant effect of attention on contrast appearance using an equality task. They claim that such equality judgments are bias-free whereas comparative judgments are bias-prone and propose an alternative interpretation of the previous findings based on a decision bias. However, to date there is no empirical support for the superiority of the equality procedure. Here, we compare biases and sensitivity to shifts in perceived contrast of both paradigms. We measured contrast appearance using both a comparative and an equality judgment. Observers judged the contrasts of two simultaneously presented stimuli, while either the contrast of one stimulus was physically incremented (Experiments 1 and 2) or exogenous attention was drawn to it (Experiments 3 and 4). We demonstrate several methodological limitations of the equality paradigm. Nevertheless, both paradigms capture shifts in PSE due to physical and perceived changes in contrast and show that attention enhances apparent contrast.

Keywords

attention; appearance; psychophysical methods; contrast perception; spatial vision

Introduction

Attention and appearance

Selective attention enhances processing of behaviorally relevant aspects of a visual scene at the cost of others. Thereby attention enables us to deal with the enormous amount of available sensory information. Attending covertly (without eye movements) to a particular location in space affects performance both when attention is allocated voluntarily— endogenous or sustained attention—or involuntarily—exogenous or transient attention (Egeth & Yantis, 1997; Posner, 1980; Yantis, 2000). Attention improves accuracy, spatial

resolution, and processing speed for stimuli at the attended location (e.g., Cameron, Tai, Eckstein, & Carrasco, 2004; Carrasco, Giordano, & McElree, 2006; Carrasco & McElree, 2001; Carrasco, Penpeci-Talgar, & Eckstein, 2000; Dobkins & Bosworth, 2001; Hawkins et al., 1990; Ling & Carrasco, 2006a; Liu, Stevens, & Carrasco, 2007; Nakayama & Mackeben, 1989; Posner, 1980; Talgar, Pelli, & Carrasco, 2004; Yeshurun & Carrasco, 1998, 1999) while it impairs performance at the unattended locations (Giordano, McElree, & Carrasco, 2009; Ling & Carrasco, 2006b; Lu & Dosher, 1998; Luck et al., 1994; Montagna, Pestilli, & Carrasco, 2009; Pestilli & Carrasco, 2005; Pestilli, Viera, & Carrasco, 2007; Posner, 1980).

Whereas these effects of attention on visual *performance* are well established, a recent series of experiments demonstrates that attention also modulates the *phenomenological appearance* of several low-level stimulus features, a proposition that has been debated for more than a century (James, 1890/1950; von Helmholtz, 1910). To test subjective perception, a paradigm has been developed in which observers perform a task contingent upon a comparative judgment between two stimuli, while an uninformative cue directs spatial attention to one of the stimuli (Carrasco, Ling, & Read, 2004). For example, observers are asked to report the orientation of the higher contrast stimulus. By pressing one key, observers convey information regarding both properties, and they implicitly report their subjective experience of contrast. Changes in apparent contrast are measured in terms of shifts of the point of subjective equality (PSE), at which the two stimuli appear equal. This 2 × 2 alternative forced choice (AFC) paradigm enables an objective and rigorous study of attention and subjective experience (Luck, 2004; Treue, 2004). Using this paradigm, attention has been shown to alter appearance of contrast (Carrasco, Fuller, & Ling, 2008; Carrasco et al., 2004; Fuller, Park, & Carrasco, 2009; Fuller, Rodriguez, & Carrasco, 2008; Hsieh, Caplovitz, & Tse, 2005; Liu, Abrams, & Carrasco, 2009; Störmer, McDonald, & Hillyard, 2009), spatial frequency (Abrams, Barbot, & Carrasco, 2010; Gobell & Carrasco, 2005), gap size (Gobell & Carrasco, 2005), color saturation (Fuller & Carrasco, 2006), size of a moving object (Anton-Erxleben, Henrich, & Treue, 2007), motion coherence (Liu, Fuller, & Carrasco, 2006), flicker rate (Montagna & Carrasco, 2006), and speed (Fuller et al., 2009; Turatto, Vescovi, & Valsecchi, 2007). An additional advantage of the 2 × 2 AFC paradigm is that it provides concurrent assessment of appearance and performance: In addition to altering appearance, attention improves performance at the cued location, indicating that attention has been successfully manipulated (Abrams et al., 2010; Anton-Erxleben et al., 2007; Carrasco et al., 2004; Fuller & Carrasco, 2006; Fuller et al., 2008; Ling & Carrasco, 2007; Liu et al., 2006, 2009).

In principle, the effect on appearance could also be consistent with a bias that leads observers to preferentially select the cued side of space. Several kinds of control experiments have ruled out such a cue, response or decision bias. First, when the order of cue and stimulus is reversed (postcue), a simple cue bias should persist. Instead, the effect disappears, supporting an attentional interpretation (Anton-Erxleben et al., 2007; Carrasco et al., 2008; Fuller et al., 2009; Gobell & Carrasco, 2005; Liu, Pestilli, & Carrasco, 2005). Second, when the cue–stimulus interval is lengthened to 500 ms, the effect of the cue disappears, consistent with the limited time scale of transient attention (Carrasco et al., 2004; Ling & Carrasco, 2007; Liu et al., 2006; Turatto et al., 2007). Third, a bias could arise because attention makes the secondary task (e.g., orientation discrimination) easier, so that observers might develop a tendency to select the attended stimulus, regardless of its apparent contrast. However, when observers are asked to report the orientation of the stimulus of lower, rather than higher, contrast, they choose the cued test stimulus less frequently. According to a bias explanation, observers should have chosen the cued stimulus more often regardless of the direction of the question. Instead, this result is consistent with a genuine effect of attention on appearance (Anton-Erxleben et al., 2007; Carrasco et al.,

2004; Fuller & Carrasco, 2006; Gobell & Carrasco, 2005; Ling & Carrasco, 2007; Liu et al., 2009; Montagna & Carrasco, 2006; Turatto et al., 2007).

Furthermore, a bias should be of similar magnitude independent of the stimulus properties and visual field location, but this is not the case. For instance, attention increases perceived saturation but it does not affect apparent hue, even though it improves performance in both cases (Fuller & Carrasco, 2006). Also, the effect of attention on apparent size varies inversely with standard stimulus size (Anton-Erxleben et al., 2007). Moreover, the effect of attention on apparent contrast is greater at the lower than at the upper visual vertical meridian (Fuller et al., 2008). These studies support the view that the effects of attention are not the result of a bias but rather an effect on phenomenological appearance.

Despite the evidence provided by these previous studies, the question of attentional modulation of appearance remains a matter of debate: an alternative interpretation proposes that decision biases occur in any experiment combining a cue with a comparative judgment such as evaluating which of two stimuli appears higher in contrast. Schneider and Komlos (2008) proposed that the decision bias is tied to the deployment of attention. They suggest that attention might simply change the stimulus' salience, so that the cued stimulus is prioritized leading observers to preferentially select it for their judgment. According to this account, the increase in salience need not be related to an increase in apparent contrast. They argue that the results from control experiments with a postcue and with a longer cue-target SOA are consistent with their salience-based account. Furthermore, they contend that in the reversed instructions control experiments, even when observers are asked to select the stimulus of lower contrast, observers could still associate the cued target with higher salience and simply invert their responses.

Schneider and Komlos (2008) have suggested that a paradigm that uses an equality judgment, in which observers are asked to report whether two stimuli are the same or different in contrast, provides a definite answer to settle the question of whether attention alters apparent contrast. They argue that in a comparative paradigm, PSE and criterion are confounded, whereas in an equality judgment the PSE is independent of criterion shifts. Using the equality task, they did not find a significant effect of attention on contrast appearance and thus concluded that the effects of attention in previous studies (Anton-Erxleben et al., 2007; Carrasco et al., 2004, 2008; Fuller & Carrasco, 2006; Fuller et al., 2008; Gobell & Carrasco, 2005; Ling & Carrasco, 2007; Liu et al., 2006; Montagna & Carrasco, 2006; Turatto et al., 2007) were due to a bias. The approach of using an equality judgment to disentangle an attentional effect on perception from a decisional account is intriguing. However, an examination of the validity and effectiveness of the equality paradigm is essential. It must be demonstrated that the equality task is indeed bias-free and at least as sensitive to shifts in perceived contrast as the comparative task.

Evaluating comparative and equality paradigms

Evaluating comparative and equality paradigms is relevant for visual psychophysics beyond the field of attention: comparative judgments have been used to study many different visual phenomena; for instance, adaptation effects on perceived speed (Hammett, Champion, Morland, & Thompson, 2005; Ledgeway & Smith, 1997; Stocker & Simoncelli, 2009), contrast (Hammett, Snowden, & Smith, 1994), and shape (Gheorghiu & Kingdom, 2007), perceived spatial frequencies at different eccentricities (Davis, Yager, & Jones, 1987) and at different contrasts and durations (Georgeson, 1985), speed perception (Smith & Edgar, 1990; Stone & Thompson, 1992), size perception (Charles, Sahraie, & McGeorge, 2007), vertical and horizontal meridian asymmetries (Fuller et al., 2008; Montaser-Kouhsari & Carrasco, 2009), illusions (Carlson, Moeller, & Anderson, 1984), cue combination (Ho, Landy, & Maloney, 2008), lightness constancy (Todd, Norman, & Mingolla, 2004), and

perception of 3D surface geometry (Ernst & Banks, 2002; Knill & Saunders, 2003). Should the equality judgment indeed give more bias-free estimates and at the same time be sufficiently sensitive to changes in appearance, previous interpretations of many results might have to be reevaluated.

Several psychophysical studies have investigated sensitivity and biases of comparative and equality tasks (e.g., Coltheart & Curthoys, 1968; Farell, 1985; Fetterman, Dreyfus, & Stubbs, 1996; Gorea, Caetta, & Sagi, 2005; Hautus & Lee, 1998; McKee, Klein, & Teller, 1985; Ratcliff & Hacker, 1981; Wichmann & Hill, 2001a, 2001b). In a comparative 2-AFC judgment, a bias for giving one over the other response, i.e. a shift in criterion, and a shift in PSE are indistinguishable without additional control experiments (e.g. Anton-Erxleben et al., 2007; Carrasco et al., 2004, 2008; Schneider & Bavelier, 2003; Schneider & Komlos, 2008; Shore, Spence, & Klein, 2001). Reliability of PSE estimation is affected by the slope of the psychometric function (McKee et al., 1985) as well as stimulus placement along the stimulus dimension of interest (McKee et al., 1985; Wichmann & Hill, 2001a).

Generally, in an equality judgment, the relation between criterion settings and the parameter estimates can be complex. The equality judgment replaces the one criterion of the comparative judgment with two criteria for "different because greater" and "different because lesser" responses. In the equality judgment employed by Schneider and Komlos (2008), the psychometric function is an approximately bell-shaped distribution of "same" responses, with the PSE corresponding to the peak of this distribution. The two criteria on each flank of the PSE determine the width and amplitude of the distribution, reflecting the general tendency for responding "same." Indeed, in an equality judgment of letter strings, instructions to the observers (e.g., "say same only if you are sure the stimuli are the same") affected not only speed and accuracy but also the overall proportion of "same" versus "different" responses (Ratcliff & Hacker, 1981). In Schneider and Komlos' (2008) paradigm, a bias for giving one over the other response would in principle change width and amplitude without affecting the location of the PSE (Schneider & Komlos, 2008). This assumption however only holds when the two criteria are symmetric. Asymmetric criterion settings can bias the location of the PSE. Empirically, the symmetry of the two criteria is often violated and should therefore be verified in each experiment (Petrov, 2009). However, in Schneider and Komlos' (2008) implementation of the equality paradigm symmetry of criteria was assumed, but the two criteria could not be independently calculated. Thus, the assumption that criterion shifts do not affect PSE estimation in their equality paradigm has yet to be tested.

Even if criteria were symmetric, a change in width and amplitude of the psychometric function likely influences the reliability of PSE estimation: If observers were strongly biased either to respond "same" or "different," the distribution would become very shallow. Such a strong bias is particularly likely to occur if the underlying proportion of "same" and "different" trials is unbalanced, as was the case in Schneider and Komlos' (2008) study, where "same" trials were presented only on ~11% of the trials. Such extreme probabilities lead to conservatism: If observers are unaware of the true prior odds, they overestimate the probability of occurrence of the rare event, so that their estimated odds are roughly equal to the cube root of the true prior odds (Green & Swets, 1966; Maloney, 2002). Thus, criterion settings could influence the sensitivity of the paradigm so that an effect of attention might be present but simply not captured. Furthermore, the assumption that the different parameters of a psychometric function are independent from each other has been directly tested and found to be unwarranted: The slope parameter of a psychometric function affects threshold estimation; specifically, a shallow slope impairs reliability of the threshold estimate (McKee et al., 1985). Differences in reliability between the comparative and equality paradigms have not been tested in previous studies.

Furthermore, a study that compared performance in judging a visual stimulus' duration using comparative and equality judgments revealed that accuracy in the equality task was significantly lower than in the comparative task, indicating greater difficulty of the equality task (Fetterman et al., 1996).

Aims of the current study

Based on the lack of an attentional effect with an equality paradigm, Schneider and Komlos (2008) proposed that attention does not alter appearance but rather biases decisions. However, three alternative explanations could account for their null finding: (1) There is no empirical support for the superiority of equality judgments over comparative judgments. It is possible that the equality paradigm in general is not bias-free or not sensitive to attentional effects on appearance or both. (2) Their study, in particular, might lack sufficient statistical power, and with increased power they could have found an effect with the equality task. (3) Their study lacks a baseline condition, which is essential to assess the effect of attention—instead, they compare perceived contrast in the attentional condition to the physical contrast, assuming that contrast perception without attention is veridical. However, this is an empirical question and the assumption of veridicality is not warranted (Carrasco et al., 2008; Treue, 2004).

The aims of the present study are twofold: First, we systematically investigate sensitivity and biases in comparative and equality judgments. We evaluate and compare the sensitivity and biases of the equality task as employed by Schneider and Komlos (2008) and the comparative task implemented by Carrasco and colleagues (2004) to changes in physical contrast (Experiments 1 and 2). Given that in these experiments the contrast differences are physically present, they allow us to evaluate strengths and weaknesses of the two tasks without a potential influence of attention on either appearance or decision processes. Given that comparative judgments have been used to study a variety of visual phenomena (e.g. Carlson et al., 1984; Charles et al., 2007; Davis et al., 1987; Ernst & Banks, 2002; Fuller et al., 2008; Georgeson, 1985; Gheorghiu & Kingdom, 2007; Hammett et al., 1994, 2005; Ho et al., 2008; Knill & Saunders, 2003; Ledgeway & Smith, 1997; Montaser-Kouhsari & Carrasco, 2009; Smith & Edgar, 1990; Stocker & Simoncelli, 2009; Stone & Thompson, 1992; Todd et al., 2004), this systematic comparison has implications beyond the question of whether attention alters appearance.

Second, we evaluate and compare the effect of attention on apparent contrast with equality and comparative tasks (Experiments 3 and 4). If the equality paradigm were bias-free and at least as sensitive as the comparative paradigm (Experiment 1) and if the equality paradigm were not to reveal significant effects of attention (Experiment 3), then these findings would challenge our interpretation of previous results regarding the effects of cueing on the appearance of contrast (Carrasco et al., 2004, 2008; Fuller et al., 2008, 2009; Hsieh et al., 2005; Ling & Carrasco, 2007; Liu et al., 2009; Störmer et al., 2009) and of other visual attributes (Abrams et al., 2010; Anton-Erxleben et al., 2007; Carrasco et al., 2008; Fuller & Carrasco, 2006; Fuller et al., 2009; Gobell & Carrasco, 2005; Liu et al., 2006; Montagna & Carrasco, 2006; Turatto et al., 2007).

Experiments 1 and 2: Physical contrast increments

In Experiments 1 and 2, observers judged the contrasts of two simultaneously presented stimuli. We incremented the contrast of one stimulus to simulate the hypothesized effect of attention by changing the physical contrast. Any method that is sensitive to a change in apparent contrast must be able to capture corresponding differences in physical contrast. We chose three different levels of contrast enhancement (3%, 6%, and 9% away from the reference contrast) to test whether the equality and comparative paradigms would differ in

the smallest effect size they could detect. The magnitude of the three contrasts increments was chosen so that the medium level roughly corresponded to the effect size reported with exogenous or transient attention (Carrasco et al., 2004, 2008; Fuller et al., 2008, 2009; Ling & Carrasco, 2007). The only difference between the two experiments was the instruction: In Experiment 1, observers were asked to report if the contrasts were the same or different; in Experiment 2, they were asked to indicate the location of the stimulus of higher contrast.

Experiments 1 and 2 were designed to assess biases and sensitivity of comparative and equality judgments such as the one used by Schneider and Komlos (2008); therefore, we kept our equality task very similar to theirs. However, in order to maximize the sensitivity of both methods, we varied stimulus contrast in finer steps than previous studies (Carrasco et al., 2004, 2008; Fuller et al., 2008, 2009; Ling & Carrasco, 2007; Schneider & Komlos, 2008).

Methods

Observers—Nine observers (4 male and 5 female; mean age = 28 years, *SD* = 5) participated in both experiments for remuneration (as well as experiments 3 and 4, see below); the order of all four experiments was randomized so that no two observers encountered the same order. All observers were naive to the purpose of the experiments and had normal or corrected-to-normal vision. Some observers were experienced and others were inexperienced in visual psychophysical tasks. The institutional review board of New York University approved all procedures.

Apparatus—Experiments were performed in a dark experimental room. Stimuli were presented on a calibrated and linearized CRT monitor (IBM P260) with a viewable area of 40×30 cm, a resolution of 1 cm/deg (corresponding to 32 pixels/deg), and a refresh rate of 106 Hz. Stimuli were presented on a medium gray background (53 cd/m^2). Observers used a chin rest positioned at a distance of 57 cm from the monitor. Experiments were run on an Apple Macintosh computer (iMac). Stimulus presentation and recording of the observers' responses was controlled by a custom MATLAB (The MathWorks, Natick, MA) script using the Psychophysics Toolbox extension (Brainard, 1997; Pelli, 1997).

Stimuli and trial sequence—Figure 1A shows the trial sequence for Experiments 1 and 2. Each trial started with the presentation of a dark (<1 cd/m^2) fixation square of 0.2 deg side length at the center of the screen. After 510 ms, a test and a standard stimulus were presented simultaneously at 4 deg eccentricity left and right of fixation for 50 ms. Within each condition, the contrast of the standard stimulus was fixed whereas the contrast of the test stimulus was varied (see Figure 1B). Stimuli were stationary 4-cpd Gabors with a Gaussian envelope of 1 deg diameter at half height. In each trial, both Gabors were tilted either 45° to the left or to the right of vertical. The phase of each Gabor was randomly varied. Both phase shifts and tilts were introduced to minimize adaptation. In the baseline condition (no increment), the standard stimulus had a Michelson contrast of 26% while the test contrast was varied in 19 steps on a logarithmic scale between 10% and 67% (Figure 1B). In six increment conditions, either the test or the standard contrast was incremented by either ~0.05, ~0.1, or ~0.15 log contrast steps (Figure 1B). For the 26% contrast standard stimulus these increments correspond to ~3%, 6%, or 9% contrast. The seven different conditions (six increments and one baseline) were interleaved and presented with the same frequency in random order within each block. In each trial, the test contrast was randomly chosen from the 19 values with equal probability.

Procedure—Experiments consisted of two 1-hour sessions performed on different days. Sessions were divided into 4 blocks of 350 trials, so that each observer completed 2800

trials. This corresponds to ~21 trials per data point (133 combinations of condition and test contrast). In the beginning of each block, observers received written instructions on the screen, asking either "Do the two stimuli have the same or different contrast?" (Experiment 1) or "Which stimulus has higher contrast?" (Experiment 2). In all experiments, a 2-AFC paradigm was used. In the equality task (Experiment 1), observers were asked to indicate if both stimuli appeared to have the same or a different contrast by pressing either the "s" or the "d" key, respectively. In the comparative task (Experiment 2), they pressed the left-arrow and right-arrow keys to indicate that the left or the right stimulus had higher contrast, respectively. Responses could be given immediately after offset of the Gabors and the fixation square; response time was unlimited.

Analysis—For the equality experiment (Experiment 1), we analyzed the proportion of "same contrast"—responses as a function of the logarithm of the physical test contrast for each increment condition. Using non-linear regression in MATLAB (The MathWorks, Natick, MA), these data were fit with a scaled Gaussian function with the parameters mean (which is equivalent to the PSE), standard deviation, and amplitude (scale factor). For the comparative experiment (Experiment 2), we analyzed for each increment condition the proportion of trials in which an observer responded "test had higher contrast" as a function of the logarithm of the physical test contrast. We fit these data with a cumulative Gaussian function with the parameters mean and standard deviation. For both models, the standard deviation parameter was constrained to positive values, while the amplitude parameter in Experiment 1 was constrained to be between 0 and 1, as the frequency of responding "same" can never be more than 100%. Data points at different contrast levels were weighted proportionally to their number of repetitions. Effects of judgment type and increment were analyzed with a repeated measures ANOVA using SPSS (SPSS Inc., IL). Whenever Mauchly's test of sphericity indicated a violation of the assumption of sphericity, we report Greenhouse–Geisser-corrected p-values; otherwise, uncorrected p-values are reported. Unless indicated otherwise, all subsequent pairwise comparisons were two-tailed.

Results

Goodness of fit—The scaled Gaussian model fit the data from the equality judgment (Experiment 1) reasonably well; average R^2 was 0.8539 ($SEM = 0.0357$). The data from the comparative experiment (Experiment 2) were well described by the cumulative Gaussian model (average $R^2 = 0.9690$, $SEM = 0.0070$). For all further analysis in Experiments 1 and 2, we excluded one observer who had R^2s < 0.7 in several conditions in the equality experiment (Experiment 1).

Estimation of PSE—Figure 2A shows the average results of the equality experiment (Experiment 1). For each increment condition, the proportion of "same" responses is plotted as a function of baseline test contrast, together with the scaled Gaussian fit to the average data. Figure 2B shows the average data for the comparative task (Experiment 2). For each increment condition, we plot the proportion of trials in which observers reported that the test stimulus had higher contrast as a function of baseline test contrast together with the cumulative Gaussian fit to the average data.

In the equality experiment, the PSE in the "no increment" (baseline) condition is at 23.6%, i.e., at a lower contrast than the true standard contrast (black curve, true standard contrast 26%). When the contrast of the standard stimulus is incremented by 0.05 (dark red), 0.1 (medium red), or 0.15 (light red) log contrast steps, the PSE increases to higher test contrasts of 26.0%, 29.3%, and 32.6%, respectively. When the contrast of the test stimulus is incremented by 0.05 (dark blue), 0.1 (medium blue), or 0.15 (light blue) log contrast steps,

the PSE decreases to lower baseline test contrasts of 21.2%, 19.0%, and 17.1%, respectively. Table 1 summarizes the parameter estimates from the Gaussian fits.

In the comparative experiment, the PSE in the "no increment" (baseline) condition corresponds more closely to the true standard contrast (black curve, PSE 25.5%). When the contrast of the standard stimulus is incremented by 0.05 (dark red), 0.1 (medium red), or 0.15 (light red) log contrast steps, the PSE increases to a test contrast of 28.4%, 31.4%, and 35.8%, respectively. When the contrast of the test stimulus is incremented by 0.05 (dark blue), 0.1 (medium blue), or 0.15 (light blue) log contrast steps, the PSE decreases to a baseline test contrast of 23.3%, 20.6%, and 18.5%, respectively.

In Figures 2C and 2D, we plot the PSEs in the increment conditions against the PSEs without increment for each observer in the equality and comparative experiment, respectively. In both experiments, the PSEs in the different increment conditions are arranged in an orderly fashion, consistent with an increase in apparent contrast with contrast increment. The figure also shows that the PSE in the no increment condition is more variable across observers in the equality experiment than in the comparative experiment (F-test, $F(8,7) = 0.075$, $p = 0.002$).

We performed a two-way repeated measures ANOVA with factors experiment and increment (2×7) on the PSE and found a main effect of experiment ($F(1,7) = 13.718$, $p = 0.008$) and of increment ($F(6,42) = 552.855$, $p < 0.001$), but no interaction ($F(6,42) < 1$). The main effect of experiment confirms that the equality method consistently yields a lower estimate for the PSE than the comparative method, which gives PSE estimates closer to the true standard contrast. Figure 3 shows the PSE estimates averaged across observers as a function of the point of objective equality (POE) in each increment condition. The POE in the standard-incremented conditions correspond to the new (incremented) standard contrasts; in the test-incremented conditions, the POE is the contrast of the test stimulus that is (after incrementing) equal to the original standard contrast. For example, if the test values are incremented by 0.05 log contrast step, the 23% contrast test is incremented to 26% and therefore corresponds to the POE. Whereas the PSEs in the comparative experiment fall along the diagonal, the PSEs in the equality experiment are systematically shifted downward, consistent with an underestimation of the true standard contrast. To test the reliability with which the two methods estimate the location of the PSE, for each experiment we performed a t-test comparing the PSE in the baseline condition to the POE. We found no significant difference between PSE and POE in the comparative experiment ($t(8)= 1.765$, $p = 0.115$, $n = 9$), but in the equality experiment, the PSE was significantly lower than the POE ($t(7) = 2.795$, $p = 0.027$, $n = 8$; as indicated above, one observer was excluded from Experiment 1 due to low goodness of fit). This is likely related to the observation that for many observers, the left tail of the distribution of "same" responses was higher than the right tail and does not reach a value lower than ~20%.

A possible explanation for this asymmetry is that observers adopt asymmetric criteria, so that they are more likely to (incorrectly) respond "same" even at large differences between test and standard contrast when the test contrast is lower than when it is higher. This is especially puzzling because in the comparative experiment using the exact same contrast levels, the same observers are perfectly able to correctly report which of the two stimuli has higher contrast: with the lowest test contrast in the no increment condition, observers made on average ~20% errors in the equality experiment, but only ~5% in the comparative experiment.

One-way repeated measures ANOVAs for each experiment revealed a significant effect of increment on the PSE in both experiments (comparative: $F(6,48) = 416.269$, $p < 0.001$, $n =$

9; equality: $F(2.135,14.947) = 305.774$, $p < 0.001$, $n = 8$). The effect of increment shows that both methods are sensitive to shifts of the PSE when the physical contrast is altered. Pairwise comparisons indicated that in both experiments the PSEs for all increments were significantly different from one another (equality experiment: all $t(7) \geq 4.884$, $p < 0.002$, alpha adjusted for 21 comparisons: 0.0024 at an overall significance level of 0.05; comparative experiment: all $t(8) \geq 4.873$, $p < 0.001$).

Standard deviation and amplitude—The standard deviation parameter of the Gaussian fits represents the width of the distribution of "same" responses in the equality experiment and the slope of the psychometric function in the comparative experiment. In both paradigms, the standard deviation is determined by the observer's ability to discriminate the different contrasts as well as his/her criterion for indicating a difference between the two stimuli. In the equality paradigm, an additional amplitude parameter was necessary. The standard deviation and amplitude parameters are both influenced by the overall proportion of "same" responses independent of actual stimulus contrast, which is dependent on the criterion: If observers are inclined to respond "same," this is reflected in higher standard deviation and/or higher amplitude. In the equality experiment, the standard deviation and amplitude parameters are clearly interdependent (see below). Note that there is no amplitude parameter in the comparative paradigm. Therefore, the comparison of the standard deviation between the two paradigms should be interpreted cautiously.

A two-way repeated measures ANOVA (2 experiments × 7 increments) on the standard deviation parameter revealed that the standard deviation was marginally lower in the comparative than the equality experiment ($F(1,7) = 4.321$, $p = 0.079$). There was a main effect of increment ($F(6,42) = 6.406$, $p < 0.001$), but the interaction of experiment and increment was not significant ($F(2.799,19.59) = 1.802$, $p = 0.182$). In both experiments, the effect of increment was confirmed in separate one-way repeated measures ANOVAs (equality: $F(6,42) = 5.194$, $p < 0.001$; comparative: $F(6,42) = 4.569$, $p = 0.001$). Whereas the ANOVA yielded significant effects, none of the pairwise comparisons reached significance after correction for multiple comparisons (all $p > 0.004$, alpha adjusted for 21 comparisons: 0.0024 at an overall significance level of 0.05).

For the equality experiment, the amplitude parameter also varied significantly with increment (one-way ANOVA; $F(6,42) = 6.837$, $p < 0.001$). There is a tendency for the standard incremented—conditions to have a lower amplitude than the test incremented—conditions, although none of the pairwise comparisons reached significance after correction for multiple comparisons (all $p > 0.004$).

Generally, in the equality experiment, as the distributions shift to the right their amplitudes and their standard deviations become lower. These changes could be related to asymmetric criteria, which might make observers more likely to judge a low contrast test stimulus as equal to the standard than a high contrast test stimulus.

Interdependence of parameter estimates—We analyzed the baseline condition without contrast increment and found not only a strong correlation between standard deviation and amplitude in the equality experiment (Spearman's rho (6) = 0.95, $p = 0.0013$), but also a trend for a negative correlation between PSE and standard deviation (Spearman's rho (6) = −0.64, $p = 0.096$) as well as for PSE and amplitude (Spearman's rho (6) = −0.70, $p = 0.065$). These findings suggests that the parameter independence in the equality judgment, as assumed by Schneider and Komlos (2008), is not justified. In contrast, there was no correlation between PSE and standard deviation in the comparative experiment (Spearman's rho (7) = −0.12, $p = 0.776$).

To further investigate influences of criterion settings on the parameters in the equality paradigm, we conducted ideal observer simulations (see Appendix). In the equality experiment, we found that if observers deviate from the ideal observer by adjusting their criteria to the extreme probabilities of the occurrence of "same" trials (Green & Swets, 1966; Maloney, 2002), standard deviations and amplitudes became larger. These values corresponded to the empirical data more closely than the ideal observer simulation. This result suggests that observers develop a bias for responding "same." Note that in all simulations, symmetric criteria were modeled; thus, unlike the real observers' PSE, the simulated PSE was not biased and was independent of the other parameters.

Reaction times—We performed a two-way repeated measures ANOVA with the factors experiment and increment (2×7) on observer response times for the perceptual judgment. Reaction times were significantly longer in the equality than in the comparative experiment (515 ms ± 23 ms *SEM* vs. 380 ms ± 29 ms *SEM;* $F(1,8) = 25.993$, $p = 0.001$). However, there was no effect of increment ($F(1.607,12.859) = 2.431$, $p = 0.134$) and no interaction ($F(1.478,11.821) < 1$). The effect of experiment on reaction time is consistent with the idea that observers find the equality task more difficult (Fetterman et al., 1996).

Experiments 3 and 4: Attention

In Experiments 3 and 4, observers judged the contrasts of two simultaneously presented stimuli, one of which appeared on the same side as a transient attentional cue. The only difference between the two experiments was the instruction: In Experiment 3, we asked observers to report if the contrasts were the same or different; in Experiment 4, we asked them to report the stimulus of higher contrast.

To compare our results with those of Schneider and Komlos (2008), we kept the equality task as similar as possible to theirs, but to improve the sensitivity of the equality task, we introduced some critical differences: First, we varied stimulus contrast in finer steps. Second, we tested not only one (attentional) condition, but three conditions in which the cue was either peripheral (standard-cued or test-cued) or central (neutral cue), which is necessary to assess the attention effect.

The comparative task (Experiment 4) was similar to the study by Carrasco et al. (2004; see also Carrasco et al., 2008; Fuller et al., 2008, 2009; Ling & Carrasco, 2007), with the difference that in their study observers had to perform an orientation discrimination task contingent on the contrast discrimination; here we did not include such a contingent task to keep the paradigm comparable to the equality method in which a 2×2 AFC design is not possible.

Methods

Observers—The same 9 observers who also participated in Experiments 1 and 2 participated in these experiments; the order of all four experiments was randomized so that no two observers encountered the same order.

Apparatus—The same apparatus was used as in Experiments 1 and 2.

Stimuli and trial sequence—Figure 1C shows the trial sequence for Experiments 3 and 4. The stimuli were identical to those used in Experiments 1 and 2. Each trial began with the presentation of the fixation square. After 510 ms, a brief (70 ms) stationary cue was flashed. The cue was a dark (<1 cd/m^2) square of 0.3 deg side length; its position was either 1.5 deg above the test position (test cued condition), 1.5 deg above the standard position (standard

cued condition), or exactly at the fixation point (neutral condition). After an interstimulus interval of 50 ms, the two Gabors were presented.

Procedure—As in Experiments 1 and 2, each experiment was conducted in two sessions of 1 hour each on different days. In Experiments 3 and 4, each of the 8 blocks contained 300 trials, yielding 2400 trials in total. This corresponds to ~42 trials per data point (57 combinations of condition and test contrast). Otherwise the procedure was the same as in Experiments 1 and 2.

Analysis—Data from each cue condition in the equality experiment (Experiment 3) were fit with a scaled Gaussian function while data from each cue condition in the comparative experiment (Experiment 4) were fit with a cumulative Gaussian function. All fitting routines and subsequent tests followed the same procedures as in Experiments 1 and 2.

Results

Goodness of Fit—The scaled Gaussian model fit the data from the equality judgment (Experiment 3) reasonably well: Average R^2 was 0.8956 ($SEM = 0.0205$). The data from the comparative experiment (Experiment 4) were well described by the cumulative Gaussian model (average $R^2 = 0.9777$; $SEM = 0.0068$).

Effects of attention on apparent contrast—Figure 4A shows the average results of the attention experiment using the equality task. For each cue condition, we plot the proportion of "same" responses as a function of baseline test contrast, together with the scaled Gaussian fit to the average data. Figure 4B shows the average results of the attention experiment using the comparative task. For each cue condition, we plot the proportion of trials in which observers reported that the test stimulus had higher contrast than the standard stimulus as a function of baseline test contrast together with the cumulative Gaussian fit to the average data. Table 2 summarizes the parameter estimates from the Gaussian fits.

In the equality experiment, the PSE in the neutral condition is at 22.6% (black curve). When attention is drawn to the location of the test stimulus (blue), the PSE shifts to a lower contrast (21.3%); when the standard stimulus is attended (red), the PSE shifts to a higher contrast (24.1%). In the comparative experiment, the PSE in the neutral cue condition is approximately at the true standard contrast (black curve, 25.3%). When attention is drawn to the location of the test stimulus (blue), the PSE shifts to a lower contrast (21.5%), whereas when the standard stimulus is attended (red), the PSE shifts to a higher contrast (28.4%).

In Figures 4C and 4D, the PSEs in the attention conditions are plotted as a function of the PSE in the neutral condition for each observer in the equality and comparative experiments. In both experiments, almost all observers' PSE in the standard cued condition are above the diagonal; that is the PSE in the standard cued condition is at higher contrast than in the neutral condition, and conversely, almost all test cued PSEs are below the diagonal, consistent with an increase in apparent contrast by attention. These figures also show that the PSE in the neutral condition is more variable across observers in the equality than the comparative experiments (*F*-test, $F(8,8) = 0.104$, $p = 0.004$).

We compared the effects of experiment type and cue condition on the PSE with a 2×3 repeated measures ANOVA. We found a main effect of experiment ($F(1,8) = 8.397$, $p = 0.020$) and of cue ($F(1.034,8.274) = 8.206$, $p = 0.020$) and a marginal interaction ($F(1.117,8.939) = 4.671$, $p = 0.056$). A separate one-way repeated measures ANOVA of the effect of the cue for each experiment yielded a significant effect in both experiments (equality: $F(1.105,8.838) = 7.88$, $p = 0.019$; comparative: $F(1.038,8.307) = 7.364$, $p = 0.025$). Paired samples *t*-tests (one-tailed, alpha adjusted for three comparisons: 0.0170 at an

overall significance level of 0.05) indicated that in the equality experiment, the PSE in the test-cued condition was significantly lower than the neutral PSE ($t(8) = 3.447$, $p = 0.005$) and significantly lower than the standard-cued PSE ($t(8) = 2.922$, $p = 0.010$), whereas the neutral PSE was marginally lower than the standard ($t(8) = 2.231$, $p = 0.028$). In the comparative experiment, the PSE in the test-cued condition was significantly lower than the neutral PSE ($t(8) = 3.361$, $p = 0.005$) and the standard-cued PSE ($t(8) = 2.730$, $p = 0.013$), whereas neutral and standard-cued PSE were not significantly different ($t(8) = 2.050$, $p = 0.038$). These shifts of the PSE in both experiments are consistent with the hypothesis that attention increases perceived contrast.

The PSE in the neutral condition is expected to be approximately equal to the physical standard contrast. As in the increment experiments, this is the case in the comparative experiment (*t*-test comparing the PSE to the POE, $t(8) = 1.894$, $p = 0.095$, Cohen's $d = 0.628$), but not in the equality experiment ($t(8) = 3.792$, $p = 0.005$, Cohen's $d = 1.263$). These results indicate that the neutral PSE estimates from the comparative paradigm are closer to the POE than those of the equality paradigm. As in the increment experiments, in the equality judgment the left tail of the distribution of "same" responses was higher than the right tail for many observers; whereas such an asymmetry was also present in the comparative judgment, it was not as pronounced as in the equality judgment.

Effects of attention on standard deviation and amplitude—We performed a 2×3 repeated measures ANOVA with factors experiment and cue condition on the standard deviation parameter. There was a trend for larger standard deviations in the equality experiment ($F(1,8) = 4.191$, $p = 0.075$), a main effect of cue condition ($F(2,16) = 13.547$, $p < 0.001$), and an interaction ($F(2,16) = 6.813$, $p = 0.008$). Pairwise comparisons (alpha adjusted for three comparisons: 0.0170 at an overall significance level of 0.05) revealed that in the equality experiment, the standard deviation parameter was neither different between the neutral and the standard-cued condition ($t(8) = 2.418$, $p = 0.042$), nor between the test-cued and standard-cued condition ($t(8) = 1.989$, $p = 0.082$), or between test-cued and neutral condition ($t(8) < 1$). In the comparative experiment, the standard deviation parameter was larger in the test-cued than in the neutral and the standard-cued conditions ($t(8) = 4.753$, $p = 0.001$ and $t(8) = 3.982$, $p = 0.004$, respectively) but was not different between the neutral and standard-cued condition ($t(8) < 1$).

In the equality experiment, attention significantly affected the amplitude parameter (one-way repeated measures ANOVA, $F(2,16) = 20.970$, $p < 0.001$): The amplitude was higher in the neutral than in both attentional conditions (paired samples *t*-tests, test cued versus neutral: $t(8) = -4.553$, $p = 0.002$, neutral versus standard cued: $t(8) = 5.301$, $p = 0.001$), meaning that the presentation of the cue decreased the overall number of "same" responses. Also, the amplitude was higher when the test was cued than when the standard was cued ($t(8) = 2.813$, $p = 0.023$). Like in the increment experiments, as the distribution shifts to the right amplitude and standard deviation become lower. This could be related to asymmetric criteria, which might make observers more likely to judge a low contrast test stimulus as equal to the standard than a high contrast test stimulus.

Reaction times—We compared the effects of experiment and cue on reaction time in a 2×3 repeated measures ANOVA. As in the increment experiments, reaction times were significantly longer in the equality than in the comparative experiment (605 ms ± 47 ms *SEM* vs. 447 ms ± 41 ms *SEM*; $F(1,8) = 30.97$, $p = 0.001$), with no effect of cue ($F(1.088,8.704) < 1$) and no interaction ($F(1.239,9.902) = 1.557$, $p = 0.248$). Again, the main effect of experiment is consistent with the notion that observers have more difficulty making equality than comparative judgments (Fetterman et al., 1996).

Evaluating sensitivity of the comparative and equality paradigm—One possible reason for the discrepancy between Schneider and Komlos' (2008) findings and our results is that the equality judgment is generally less sensitive than the comparative judgment (see also Experiments 1 and 2). To further investigate the sensitivity of both methods to attentional effects on apparent contrast, we evaluated the statistical power of both the comparative and equality paradigms. We repeatedly subsampled the data from Experiments 3 and 4 and tested how frequently each paradigm would reveal a significant effect of attention on apparent contrast. We included half of the trials at 9 of the original contrast levels (centered on the 26% contrast of the standard stimulus, skipping every other value), so that we encompassed a similar contrast range (11 to 60%), and used the same number of contrast levels and trials per test contrast level and condition (~20) as Schneider and Komlos (2008) did. (Note that they tested 5 more observers than the present study). We subsampled the observers' responses by randomly drawing half of the trials completed by the 9 observers and we ran 1000 repetitions. For each repetition of a given experiment, we analyzed the data as described above and discarded observers whose goodness-of-fit (R^2) was <0.7 in one or more of the cue conditions. Unless this eliminated 4 or more observers, we then subjected the PSEs derived from these subsampled data to a one-way ANOVA. With the comparative paradigm, the cue had a significant ($p \leq 0.05$) effect on the PSE in 986 out of 1000 repetitions (98.6%), whereas with the equality paradigm, the effect was significant only in 307 out of 947 repetitions (32.4%). Although this analysis cannot provide direct evidence for lower sensitivity of Schneider and Komlos' (2008) particular implementation of the equality paradigm, this pronounced difference suggests that in general, the equality paradigm is less sensitive than the comparative paradigm. Therefore, it is possible that Schneider and Komlos (2008) might have missed the effect of attention with the equality judgment due to the low sensitivity of the equality paradigm.

Discussion

We measured the PSE in a comparative and an equality 2-AFC contrast judgment and found that both paradigms are able to capture shifts in PSE due to changes in physical contrast (Experiments 1 and 2), as well as changes in apparent contrast induced by an attentional cue (Experiments 3 and 4).

When the contrast of one of two stimuli was incremented by three different levels, the PSE shifted accordingly in both paradigms (Experiments 1 and 2). In addition, we observed effects on the shape of the distribution of responses for the equality and comparative tasks. PSE estimates derived from the equality judgment are consistently underestimated relative to veridical contrast. Furthermore, PSE estimates were more variable across observers in the equality task than in the comparative task.

Allocating exogenous attention to one stimulus led to a shift in PSE consistent with an increase in apparent contrast of the attended stimulus (Experiments 3 and 4). In the comparative judgment, the increase in apparent contrast was ~3.5% relative to the neutral condition, and ~7% relative to the unattended stimulus. Using the equality task, we observed similar but smaller changes in PSE of ~1.4% contrast relative to the neutral condition, and of ~2.8% relative to the unattended stimulus. Both the paradigm and the attentional cue affected the shape of the distribution. Again, PSE estimates in the neutral condition (without focal attention) were more variable across observers in the equality task than in the comparative task.

Comparison of sensitivity and biases in the comparative and equality paradigms

The comparison of the comparative and equality paradigms has implications for visual psychophysics beyond the study of contrast appearance. For example, comparative

paradigms have been used to investigate adaptation (Gheorghiu & Kingdom, 2007; Hammett et al., 1994, 2005; Ledgeway & Smith, 1997; Stocker & Simoncelli, 2009), perception of spatial frequency (Davis et al., 1987; Georgeson, 1985), speed (Smith & Edgar, 1990; Stone & Thompson, 1992), and size (Charles et al., 2007), vertical and horizontal meridian asymmetries (Fuller et al., 2008; Montaser-Kouhsari & Carrasco, 2009), illusions (Carlson et al., 1984), cue combination (Ho et al., 2008), lightness constancy (Todd et al., 2004), and perception of 3D surface geometry (Ernst & Banks, 2002; Knill & Saunders, 2003). Equality paradigms have been used for example to study temporal order perception (Heron, Hanson, & Whitaker, 2009; Schneider & Bavelier, 2003; Van der Burg, Olivers, Bronkhorst, & Theeuwes, 2008). Hence, it is essential to evaluate sensitivity and biases of these paradigms without the possible interaction of attention.

In comparative paradigms on appearance, a shift in PSE and a shift in criterion, for example a tendency to select the cued rather than the uncued stimulus, or the lower rather than the upper visual field location, are indistinguishable without additional control experiments (e.g. Anton-Erxleben et al., 2007; Carrasco et al., 2004, 2008; Schneider & Bavelier, 2003; Schneider & Komlos, 2008; Shore et al., 2001). However, in Experiments 1 and 2 we demonstrate several methodological limitations of the equality task, in line with previous research (Farell, 1985; Fetterman et al., 1996; Gorea et al., 2005; Hautus & Lee, 1998; Ratcliff & Hacker, 1981). The equality paradigm is less sensitive and is not exempt from biases that complicate the interpretation of PSE shifts.

First, for the equality task, three instead of two parameters are necessary to explain the data. Despite the additional free parameter, the Gaussian model provides lower goodness of fit for the equality data than for the comparative data. The scaling parameter is needed because when the two stimulus contrasts are physically the same, observers respond "same" only in ~80% of the trials. Counterintuitively, this corresponds to a bias for responding "same," not different, as is clarified by the ideal observer model: Given that "same" and "different" trials were imbalanced, so that "same" trials occurred only with ~5% probability, an ideal observer with optimal and symmetric criteria would respond "same" only in ~65% of the trials in which the stimulus contrasts actually are the same. If observers adjusted criteria and overestimated the probability of "same" trials (Green & Swets, 1966; Maloney, 2002), the observer model would predict that they respond "same" in ~77% of actual "same" trials. This simulation result is closer to the proportion of "same" responses in our empirical data. Additionally, we found a trend for wider distributions (larger standard deviations) than in the comparative judgment, which is consistent with a criterion that favors "same" responses.

According to Schneider and Komlos (2008), these biases do not affect PSE estimation. However, empirically, modulation of amplitude and standard deviation may lead to very shallow distributions of "same" responses, which renders the measurement of the PSE parameter less reliable. Schneider and Komlos (2008) weighted their observers' parameter estimates by the reliability of those estimates. However, they did not evaluate whether the estimates from the equality judgment were as reliable as those from their comparative judgment. We assessed this issue in our data and found that PSE estimates are more variable across observers in the equality than in the comparative judgment. For instance, in the control condition in which neither of the contrasts was incremented, standard error across observers was approximately four times larger in the equality experiment than in the comparative experiment.

Furthermore, on average, the distribution of "same" responses is not symmetric around the point of physical equality. Many observers are more likely to respond "same" when the test contrast is lower than the standard contrast than when it is higher. It is unlikely that this bias is due to a difference in sensitivity (discrimination ability) for low and high contrasts,

because the same observers have no difficulty identifying the difference in test and standard contrast for the exact same contrast values in the comparative experiment. Instead, it is likely that observers adopt different criteria for low and high contrasts. This finding is consistent with research demonstrating that the assumption of symmetric criteria in an equality task is often violated and that experimenters should test symmetry for each experimental condition (Petrov, 2009). In experimental designs that are completely counterbalanced, so that at each stimulus level "same" and "different" trials occur with equal probability, criteria can be independently calculated. Due to the high number of stimulus levels and because we wanted to keep our paradigm as similar as possible to Schneider and Komlos' (2008) study, we did not use such a design. Therefore, in the present study, as well as in Schneider and Komlos (2008), it is impossible to calculate criteria without the assumption of symmetry because the equality judgment discards information about the direction of the perceived difference. The effects of increment on the standard deviation and the amplitude parameter in the equality paradigm may be related to asymmetric criteria: The asymmetry might be exacerbated as the distribution moves toward lower test contrasts, leading to higher amplitudes and larger standard deviations. More importantly, the asymmetry in the data can bias the location of the PSE when symmetric criteria are (inappropriately) assumed. Consistent with such a bias, there were trends for a correlation between PSE and standard deviation as well as for PSE and amplitude parameters in the equality experiment. This is in line with the finding that threshold estimation is not independent from the shape of the psychometric function (McKee et al., 1985). In contrast, in the comparative experiment, there was no correlation between PSE and standard deviation.

Only one other study (Valsecchi, Vescovi, & Turatto, 2010) has attempted to test the sensitivity of an equality judgment such as the one implemented by Schneider and Komlos (2008). They report an effect of attention on apparent speed with a comparative but not with an equality judgment and conclude that attention may affect the salience but not the appearance of the attended stimulus. However, their equality paradigm suffers from the same issues we have described: (1) amplitude, standard deviation, and PSE are all significantly correlated with one another indicating that the parameters are not independent (Valsecchi et al., 2010; Experiment 4); (2) the PSE estimates (without attention) are clearly different from the POE in their equality experiment; (3) their data in the equality paradigm are asymmetric; and (4) the fits to the data from the equality paradigm are consistently poorer than those from the comparative paradigm. Schneider and Komlos's (2008) main argument for the superiority of equality over comparative judgments is that in the equality judgments the PSE is unbiased. However, this assertion is not warranted because, in Valsecchi et al.'s (2010) as well as in the present study, the PSE in the equality judgment is not independent from the other parameters and is affected by asymmetries. Furthermore, Valsecchi et al.'s (2010) data provide converging evidence that equality judgments are less sensitive and that veridical perception without attention cannot be assumed. These findings are problematic for the interpretation of Schneider and Komlos's (2008) null results with the equality judgment.

Reaction times were longer in the equality than in the comparative judgment by >100 ms, indicating that the equality judgment was more difficult. This is consistent with observers' reports that they subjectively perceived the equality task to be harder than the comparative judgment. Correspondingly, previous research shows that performance in the equality task is lower than in the comparative task (Fetterman et al., 1996; Turatto et al., 2007). Similarly, in Schneider and Komlos (2008) study, response times were slower in the equality judgment than in the comparative judgment. We propose that that finding is consistent with a difference in task difficulty. One could argue that an additional cognitive step is required with the equality task—observers make the easier comparative judgment first and then

remap their decision to the equality judgment. Thus, the higher difficulty may contribute to the noisier estimates in the equality experiment.

Schneider and Komlos (2008) suggest that equality judgments are not prone to response biases and are therefore the superior method to study appearance. Conversely, using a design that avoids potential biases introduced by attentional cues, we demonstrate that the equality judgment is less sensitive and suffers from other biases, which are related to criterion shifts and criterion asymmetries that cannot be measured. These methodological issues with the equality paradigm have implications beyond the study of contrast appearance. For example, in the study of the prior entry effect, results from equality and comparative judgments have led to different conclusions. Whereas experiments using a comparative task and exogenous attention found that an attended stimulus appears to occur earlier (Shore et al., 2001; for an alternative interpretation, see Schneider & Bavelier, 2003), with endogenous attention this effect is only found in a comparative but not in an equality judgment (Schneider & Bavelier, 2003). It has been suggested that the comparative judgment reflects both attention and an additional bias whereas the equality judgment eliminates the bias (Van der Burg et al., 2008). However, it is possible that, as in the present study, the equality paradigm was simply less sensitive than the comparative paradigm.

Attention alters appearance

Despite the lower sensitivity of equality judgments, we find a significant effect of attention on perceived contrast with both paradigms. This result differs from that of Schneider and Komlos (2008). The discrepancy could be due to one or both of the following factors: (1) lower sensitivity and statistical power of the equality paradigm and (2) the lack of a baseline condition in their study. The methodological limitations of the equality task discussed above could contribute to the lower sensitivity. Higher noise as well as a smaller magnitude of the attentional effect would result in a lower signal-to-noise ratio for the equality paradigm. In the present study, we used more contrast levels (19 instead of 9) and collected more trials for each data point (~42 instead of 20 trials), increasing the statistical power of our paradigm. Most importantly, we included a neutral condition in which none of the stimuli were focally attended and defined the attentional effect with respect to this control condition. In contrast, instead of including such a control, Schneider and Komlos (2008) assumed veridicality for a contrast judgment without attention. The present data show that this assumption is not justified: in the equality task, apparent contrasts are consistently underestimated. Due to the lack of a control condition, Schneider and Komlos (2008) could not empirically assess if the PSE was veridical without attention. Had it been underestimated (like in the present study), their data with the attention condition could actually have been consistent with an increase of apparent contrast.

In comparison to earlier studies using the comparative paradigm, the effect of attention on apparent contrast relative to the neutral condition is somewhat smaller in the present study (~3.5% compared to ~6% in Carrasco et al., 2008; Carrasco et al., 2004; Ling & Carrasco, 2007; Fuller et al., 2009). One possible reason for this is that we did not use the original 2 × 2 AFC paradigm in which observers had to perform an orientation discrimination task contingent on the contrast judgment. This paradigm was developed to minimize observers' tendency to distribute attention evenly across both the cued and the uncued locations and to obscure the purpose of the experiment. Here, the single task design might have made it easier for observers to distribute their attention.

Alternative interpretations of the effect of exogenous attention on appearance have been suggested. Schneider (2006) proposed that sensory interactions between cue and target could yield a similar shift of the PSE as an attentional increase of apparent contrast, and predicted a reversal of effects with cue contrast polarity—light cues should lead to a decrease in

apparent contrast. However, this prediction was not confirmed: Attentional cues increase apparent contrast regardless of their polarity and the magnitude of the effect is the same (Ling & Carrasco, 2007). Furthermore, using cross-modal (auditory) cues, a recent study found the same increase in apparent contrast, thus ruling out visual sensory contamination (Störmer et al., 2009).

Several types of cue biases that could account for the reported shifts in PSE have been ruled out (Carrasco, 2009). For instance, the spatial cue could invoke a tendency to simply press the key on the same side of space as the cue. This concern has been eliminated by control experiments using a postcue (Anton-Erxleben et al., 2007; Carrasco et al., 2008; Fuller et al., 2009; Gobell & Carrasco, 2005). Others have argued that if the stimuli are near the threshold of visibility, the cue itself could be mistaken for the target and so lead to a different form of cue bias (Prinzmetal, Long, & Leonhardt, 2008). However, the same effect of attention is found when all the stimuli are suprathreshold (Carrasco et al., 2008; Carrasco et al., 2004 (Experiment 2); Fuller et al., 2008; Ling & Carrasco, 2007), arguing against a confusion of target and cue due to low target visibility. Another type of cue bias could arise due to attentional modulation of performance at the cued location: Observers could be biased to select the stimulus that is easier to judge. This would predict that the effect of the cue reverses with reversing the instruction, i.e., observers would select the stimulus of lower contrast. Although some studies find a weaker effect with reversed instructions (e.g., Anton-Erxleben et al., 2007), others do not (e.g., Fuller & Carrasco, 2006; Liu et al., 2009), and a reversal has never been found (Anton-Erxleben et al., 2007; Carrasco et al., 2004; Fuller & Carrasco, 2006; Gobell & Carrasco, 2005; Liu et al., 2009; Montagna & Carrasco, 2006; Turatto et al., 2007), showing that a cue bias could at best only partially account for the effect of the cue. Furthermore, in the present study as well as in previous studies, the effect of attention occurs even if observers do not perform the concurrent discrimination at all (Anton-Erxleben et al., 2007; Carrasco et al., 2004; Schneider & Komlos, 2008).

Some authors question the value of this particular control experiment (Schneider & Komlos, 2008), arguing that even when observers are asked to select the stimulus of lower rather than higher contrast, observers could still associate the cued target with higher salience and simply invert their responses. This argument is not parsimonious because whereas such a strategy is possible, it would be unnecessarily complex from an observer's point of view. Moreover, if such an inversion took place, one would expect longer reaction times (RTs) for selecting the lower contrast stimulus than for selecting the higher contrast stimulus, but observers' RTs for the two instructions do not differ (Liu et al., 2009). A study of the prior entry effect however demonstrates that the reversed instruction control can successfully detect a bias: Whereas observers choose the cued stimulus as appearing earlier in the original experiment, they do not choose the cued stimulus as appearing later in the control experiment. Had they simply inverted their response, they would have been equally likely to choose the cued stimulus in the original as the uncued stimulus in the control experiment (Shore et al., 2001).

Schneider and Komlos (2008) argue that a cue bias can arise because attention modulates saliency, so that observers *associate* the cued target with higher contrast without actually *perceiving* it as higher in contrast. It is not clear how this hypothesis generates predictions that can be used to distinguish between salience effects and effects on contrast appearance. Moreover, effects of attention on appearance have been reported in dimensions without clear directionality with respect to salience. For example, the relation between salience in the contrast and the spatial and temporal frequency domains is not monotonic (Georgeson, 1985; Robson, 1966). Indeed, higher spatial frequencies, such as the ones used in the studies of attention and perceived spatial frequency, appear lower in contrast (Abrams et al., 2010; Gobell & Carrasco, 2005). Similarly, the increase in perceived flicker rate with attention

occurred in a range in which an increased flicker rate does not correspond to enhanced temporal contrast sensitivity (Montagna & Carrasco, 2006). These findings provide evidence against the saliency-based response bias account. As described in the Introduction, previous studies show that attentional modulation of apparent contrast as well as other stimulus features varies with stimulus dimension and specific properties, visual field location, and cue contrast (Anton-Erxleben et al., 2007; Fuller & Carrasco, 2006; Fuller et al., 2008, 2009). It is not clear how a salience-based response bias interpretation could explain either these results or the finding that attention decreases perceived brightness contrast (Tsal, Shalev, Zakay, & Lubow, 1994).

Schneider and Komlos (2008) stated that PSE estimation in the equality judgment is not prone to response biases and therefore can resolve the controversy. Our results show on the contrary that the equality judgment is not the superior method: it is less sensitive and also prone to biases. These biases are reflected in changes in amplitude and standard deviation. If criteria are symmetric, this would scale the psychometric function up or down without directly influencing the location of the PSE, although the reliability of the PSE estimation could be affected. If criteria are asymmetric, the location of the PSE can be biased.

Because in Experiment 3 standard deviation and amplitude change with the cue condition, it is important to evaluate if the attentional cue introduces this bias and the PSE shift we find in the equality paradigm could be accounted for by this change in criteria. Experiment 1 showed a (marginal) correlation between PSE and amplitude/standard deviation (without attention), but it is impossible to conclude from the correlation if the PSE location causes the change in criteria, or if asymmetric criteria cause the shift of the PSE. It is possible that attention changes perceived contrast and therefore the "test cued" curve is covering a lower (perceived) contrast range than the "standard cued" curve, and therefore the two curves are differentially affected by the criterion difference between low and high contrasts. One argument for this explanation is that the asymmetry exists without the attentional cue (in Experiment 1), and there is no reason to assume that the source of the asymmetry differs between experiments. In fact, in Experiment 1 the amplitude and standard deviation change as the curve shifts along the contrast axis. Because in this experiment there is no attentional cue, this effect on criteria can only be explained with the difference in contrast range covered.

A more prominent effect of cue condition on the amplitude parameter is that the peripheral cue led observers to respond "different" more often than in the neutral condition, resulting in lower amplitude. It is possible that although observers were instructed to base their judgment strictly on the stimulus contrast, the mere presence of the cue makes the two sides of the display different and may have led observers to report "different" more often. Such an effect of irrelevant information orthogonal to the task at hand has been observed in reaction time experiments (Dixon & Just, 1978; Eriksen & Eriksen, 1974). Another explanation could be that even though we measured responses to a fine mesh of test contrasts, the true PSE might have fallen in between two data points. This would only be problematic for the equality paradigm: Because the peak of the Gaussian distribution is not actually measured, uncertainty about its location would increase, whereas any function that is strictly monotonically increasing at that location, including any sigmoidal psychometric function, would not be affected.

In summary, although interpretation of the results of the equality experiment seems to be limited by several methodological issues, we nevertheless detect an effect of attention with an equality judgment and conclude that the null result reported by Schneider and Komlos (2008) is most likely due to the issues explained above.

The effects of attention on contrast appearance can be understood from a physiological perspective: Single cell recording studies show that an attentional effect on neuronal activity can typically be well described by a shift of the contrast-response function toward lower contrasts, consistent with the idea that attention enhances the effective contrast of a stimulus (Martinez-Trujillo & Treue, 2002; Reynolds, Pasternak, & Desimone, 2000; but see Williford & Maunsell, 2006). This enhancement of effective contrast is likely correlated with an increase in contrast *sensitivity* but also predicts the increase of contrast *appearance* with attention. Thus, physiological and perceptual effects of attention on contrast perception are expected to be strongly linked (Carrasco, 2006; Luck, 2004; Reynolds & Chelazzi, 2004; Treue, 2004). A recent study used EEG to concurrently measure the physiological correlates and behavioral effects of attention on contrast appearance (Störmer et al., 2009). They used a similar paradigm to the original one of Carrasco and colleagues (2004) but used auditory instead of visual attentional cues. They found a modulation of evoked potentials in contralateral visual cortex that correlated with the behavioral report of an increase in perceived contrast. The temporal dynamics and source location of this modulation were consistent with a boost of early sensory processing, but not with post-perceptual processes such as decision-making (Störmer et al., 2009).

Conclusion

We compared the sensitivity of equality and comparative judgments of perceived contrast with regard to physical contrast differences (Experiments 1 and 2) and attentional modulation (Experiments 3 and 4). Previous research has assumed equal sensitivity of both judgments and an absence of bias from the equality judgment. However, the present study demonstrates several methodological limitations of the equality paradigm, which may contribute to decrease the reliability of PSE estimation and render the equality judgment less sensitive to shifts in perceived contrast. Notwithstanding these methodological limitations, in this study both paradigms revealed that attention enhances apparent contrast.

Acknowledgments

References

Abrams J, Barbot A, Carrasco M. Voluntary attention increases perceived spatial frequency. Attention, Perception & Psychophysics. 2010; 72:1510–1521.

Anton-Erxleben, K.; Henrich, C.; Treue, S. Attention changes perceived size of moving visual patterns; Journal of Vision. 2007. p. 5p. 1-9.http://www.journalofvision.org/content/7/11/5

Brainard DH. The Psychophysics Toolbox. Spatial Vision. 1997; 10:433–436. [PubMed: 9176952]

Cameron EL, Tai JC, Eckstein MP, Carrasco M. Signal detection theory applied to three visual search tasks-identification, yes/no detection and localization. Spatial Vision. 2004; 17:295–325. [PubMed: 15559107]

Carlson CR, Moeller JR, Anderson CH. Visual illusions without low spatial frequencies. Vision Research. 1984; 24:1407–1413. [PubMed: 6523761]

Carrasco M. Covert attention increases contrast sensitivity: Psychophysical, neurophysiological and neuroimaging studies. Progress in Brain Research. 2006; 154:33–70. [PubMed: 17010702]

Carrasco, M. Attention: Psychophysical approaches. In: Bayne, T.; Cleeremans, A.; Wilken, P., editors. The Oxford companion to consciousness. Oxford: Oxford University Press; 2009.

Carrasco M, Fuller S, Ling S. Transient attention does increase perceived contrast of supra-threshold stimuli: A reply to Prinzmetal, Long, and Leonhardt (2008). Perception & Psychophysics. 2008; 70:1151–1164. [PubMed: 18979688]

Carrasco M, Giordano AM, McElree B. Attention speeds processing across eccentricity: Feature and conjunction searches. Vision Research. 2006; 46:2028–2040. [PubMed: 16481020]

Carrasco M, Ling S, Read S. Attention alters appearance. Nature Neuroscience. 2004; 7:308–313.

Carrasco M, McElree B. Covert attention accelerates the rate of visual information processing. Proceedings of the National Academy of Sciences of the United States of America. 2001; 98:5363–5367. [PubMed: 11309485]

Carrasco M, Penpeci-Talgar C, Eckstein M. Spatial covert attention increases contrast sensitivity across the CSF: Support for signal enhancement. Vision Research. 2000; 40:1203–1215. [PubMed: 10788636]

Charles J, Sahraie A, McGeorge P. Hemi-spatial asymmetries in judgment of stimulus size. Perception & Psychophysics. 2007; 69:687–698. [PubMed: 17929692]

Coltheart M, Curthoys I. Short-term recognition memory for pitch: Effect of a priori probability on response times and error rates. Perception & Psychophysics. 1968; 4:85–89.

Davis ET, Yager D, Jones BJ. Comparison of perceived spatial frequency between the fovea and the periphery. Journal of the Optical Society of America A. 1987; 4:1606–1611.

Dixon P, Just MA. Normalization of irrelevant dimensions in stimulus comparisons. Journal of Experimental Psychology: Human Perception and Performance. 1978; 4:36–46. [PubMed: 627849]

Dobkins KR, Bosworth RG. Effects of set-size and selective spatial attention on motion processing. Vision Research. 2001; 41:1501–1517. [PubMed: 11343718]

Egeth HE, Yantis S. Visual attention: Control, representation, and time course. Annual Review of Psychology. 1997; 48:269–297.

Eriksen BA, Eriksen CW. Effects of noise letters upon the identification of a target letter in a nonsearch task. Perception & Psychophysics. 1974; 16:143–149.

Ernst MO, Banks MS. Humans integrate visual and haptic information in a statistically optimal fashion. Nature. 2002; 415:429–433. [PubMed: 11807554]

Farell B. Same–different judgments: A review of current controversies in perceptual comparisons. Psychological Bulletin. 1985; 98:419–456. [PubMed: 4080894]

Fetterman JG, Dreyfus LR, Stubbs DA. Judging relative duration: The role of rule and instructional variables. Journal of Experimental Psychology: Animal Behavior Processes. 1996; 22:350–361. [PubMed: 8691163]

Fuller S, Carrasco M. Exogenous attention and color perception: Performance and appearance of saturation and hue. Vision Research. 2006; 46:4032–4047. [PubMed: 16979690]

Fuller S, Park Y, Carrasco M. Cue contrast modulates the effects of exogenous attention on appearance. Vision Research. 2009; 49:1825–1837. [PubMed: 19393260]

Fuller, S.; Rodriguez, RZ.; Carrasco, M. Apparent contrast differs across the vertical meridian: Visual and attentional factors; Journal of Vision. 2008. p. 16p. 11-16.http://www.journalofvision.org/content/8/1/16

Georgeson MA. Apparent spatial frequency and contrast of gratings: Separate effects of contrast and duration. Vision Research. 1985; 25:1721–1727. [PubMed: 3832596]

Gheorghiu E, Kingdom FA. The spatial feature underlying the shape-frequency and shape-amplitude after-effects. Vision Research. 2007; 47:834–844. [PubMed: 17292437]

Giordano, AM.; McElree, B.; Carrasco, M. On the automaticity and flexibility of covert attention: A speed-accuracy trade-off analysis; Journal of Vision. 2009. p. 30p. 1-10.http://www.journalofvision.org/content/9/3/30

Gobell J, Carrasco M. Attention alters the appearance of spatial frequency and gap size. Psychological Science. 2005; 16:644–651. [PubMed: 16102068]

Gorea A, Caetta F, Sagi D. Criteria interactions across visual attributes. Vision Research. 2005; 45:2523–2532. [PubMed: 15950255]

Green, DM.; Swets, JA. Signal detection theory and psychophysics. New York: John Wiley and Sons; 1966.

Hammett ST, Champion RA, Morland AB, Thompson PG. A ratio model of perceived speed in the human visual system. Proceedings of the Royal Society of London B: Biological Sciences. 2005; 272:2351–2356.

Hammett ST, Snowden RJ, Smith AT. Perceived contrast as a function of adaptation duration. Vision Research. 1994; 34:31–40. [PubMed: 8116266]

Hautus MJ, Lee AJ. The dispersions of estimates of sensitivity obtained from four psychophysical procedures: Implications for experimental design. Perception & Psychophysics. 1998; 60:638–649. [PubMed: 9628995]

Hawkins HL, Hillyard SA, Luck SJ, Mouloua M, Downing CJ, Woodward DP. Visual attention modulates signal detectability. Journal of Experimental Psychology: Human Perception and Performance. 1990; 16:802–811. [PubMed: 2148593]

Heron J, Hanson JV, Whitaker D. Effect before cause: Supramodal recalibration of sensorimotor timing. PLoS One. 2009; 4:e7681. [PubMed: 19890383]

Ho YX, Landy MS, Maloney LT. Conjoint measurement of gloss and surface texture. Psychological Science. 2008; 19:196–204. [PubMed: 18271869]

Hsieh PJ, Caplovitz GP, Tse PU. Illusory rebound motion and the motion continuity heuristic. Vision Research. 2005; 45:2972–2985. [PubMed: 15876447]

James, W. The principles of psychology. New York: Holt; 1890/1950.

Knill DC, Saunders JA. Do humans optimally integrate stereo and texture information for judgments of surface slant? Vision Research. 2003; 43:2539–2558. [PubMed: 13129541]

Ledgeway T, Smith AT. Changes in perceived speed following adaptation to first-order and second-order motion. Vision Research. 1997; 37:215–224. [PubMed: 9068821]

Ling S, Carrasco M. Sustained and transient covert attention enhance the signal via different contrast response functions. Vision Research. 2006a; 46:1210–1220. [PubMed: 16005931]

Ling S, Carrasco M. When sustained attention impairs perception. Nature Neuroscience. 2006b; 9:1243–1245.

Ling S, Carrasco M. Transient covert attention does alter appearance: A reply to Schneider (2006). Perception & Psychophysics. 2007; 69:1051–1058. [PubMed: 18018987]

Liu T, Abrams J, Carrasco M. Voluntary attention enhances contrast appearance. Psychological Science. 2009; 20:354–362. [PubMed: 19254239]

Liu T, Fuller S, Carrasco M. Attention alters the appearance of motion coherence. Psychonomic Bulletin & Review. 2006; 13:1091–1096. [PubMed: 17484441]

Liu T, Pestilli F, Carrasco M. Transient attention enhances perceptual performance and FMRI response in human visual cortex. Neuron. 2005; 45:469–477. [PubMed: 15694332]

Liu T, Stevens ST, Carrasco M. Comparing the time course and efficacy of spatial and feature-based attention. Vision Research. 2007; 47:108–113. [PubMed: 17087987]

Lu ZL, Dosher BA. External noise distinguishes attention mechanisms. Vision Research. 1998; 38:1183–1198. [PubMed: 9666987]

Luck SJ. Understanding awareness: One step closer. Nature Neuroscience. 2004; 7:208–209.

Luck SJ, Hillyard SA, Mouloua M, Woldorff MG, Clark VP, Hawkins HL. Effects of spatial cuing on luminance detectability: Psychophysical and electrophysiological evidence for early selection. Journal of Experimental Psychology: Human Perception and Performance. 1994; 20:887–904. [PubMed: 8083642]

Maloney, LT. Statistical decision theory and biological vision. In: Heyer, D.; Mausfeld, R., editors. Perception and the physical world: Psychological and philosophical issues in perception. New York: John Wiley and Sons; 2002. p. 145-189.

Martinez-Trujillo J, Treue S. Attentional modulation strength in cortical area MT depends on stimulus contrast. Neuron. 2002; 35:365–370. [PubMed: 12160753]

McKee SP, Klein SA, Teller DY. Statistical properties of forced-choice psychometric functions: Implications of probit analysis. Perception & Psychophysics. 1985; 37:286–298. [PubMed: 4034345]

Montagna, B.; Carrasco, M. Transient covert attention and the perceived rate of flicker; Journal of Vision. 2006. p. 8p. 955-965.http://www.journalofvision.org/content/6/9/8

Montagna B, Pestilli F, Carrasco M. Attention trades off spatial acuity. Vision Research. 2009; 49:735–745. [PubMed: 19385088]

Montaser-Kouhsari L, Carrasco M. Perceptual asymmetries are preserved in short-term memory tasks. Attention, Perception & Psychophysics. 2009; 71:1782–1792.

Nakayama K, Mackeben M. Sustained and transient components of focal visual attention. Vision Research. 1989; 29:1631–1647. [PubMed: 2635486]

Pelli DG. The VideoToolbox software for visual psychophysics: Transforming numbers into movies. Spatial Vision. 1997; 10:437–442. [PubMed: 9176953]

Pestilli F, Carrasco M. Attention enhances contrast sensitivity at cued and impairs it at uncued locations. Vision Research. 2005; 45:1867–1875. [PubMed: 15797776]

Pestilli, F.; Viera, G.; Carrasco, M. How do attention and adaptation affect contrast sensitivity?; Journal of Vision. 2007. p. 9p. 1-12.http://www.journalofvision.org/content/7/7/9

Petrov AA. Symmetry-based methodology for decision-rule identification in same-different experiments. Psychonomic Bulletin & Review. 2009; 16:1011–1025. [PubMed: 19966250]

Posner MI. Orienting of attention. Quarterly Journal of Experimental Psychology. 1980; 32:3–25. [PubMed: 7367577]

Prinzmetal W, Long V, Leonhardt J. Involuntary attention and brightness contrast. Perception & Psychophysics. 2008; 70:1139–1150. [PubMed: 18927000]

Ratcliff R, Hacker MJ. Speed and accuracy of same and different responses in perceptual matching. Perception & Psychophysics. 1981; 30:303–307. [PubMed: 7322806]

Reynolds JH, Chelazzi L. Attentional modulation of visual processing. Annual Review of Neuroscience. 2004; 27:611–647.

Reynolds JH, Pasternak T, Desimone R. Attention increases sensitivity of V4 neurons. Neuron. 2000; 26:703–714. [PubMed: 10896165]

Robson JG. Spatial and temporal contrast sensitivity functions of the visual system. Journal of the Optical Society of America A. 1966; 56:1141–1142.

Schneider KA. Does attention alter appearance? Perception & Psychophysics. 2006; 68:800–814. [PubMed: 17076348]

Schneider KA, Bavelier D. Components of visual prior entry. Cognitive Psychology. 2003; 47:333–366. [PubMed: 14642288]

Schneider, KA.; Komlos, M. Attention biases decisions but does not alter appearance; Journal of Vision. 2008. p. 3p. 1-10.http://www.journalofvision.org/content/8/15/3

Shore DI, Spence C, Klein RM. Visual prior entry. Psychological Science. 2001; 12:205–212. [PubMed: 11437302]

Smith AT, Edgar GK. The influence of spatial frequency on perceived temporal frequency and perceived speed. Vision Research. 1990; 30:1467–1474. [PubMed: 2247956]

Stocker, AA.; Simoncelli, EP. Visual motion aftereffects arise from a cascade of two isomorphic adaptation mechanisms; Journal of Vision. 2009. p. 9p. 1-14.http://www.journalofvision.org/content/9/9/9

Stone LS, Thompson P. Human speed perception is contrast dependent. Vision Research. 1992; 32:1535–1549. [PubMed: 1455726]

Störmer VS, McDonald JJ, Hillyard SA. Cross-modal cueing of attention alters appearance and early cortical processing of visual stimuli. Proceedings of the National Academy of Sciences of the United States of America. 2009; 106:22456–22461. [PubMed: 20007778]

Talgar, CP.; Pelli, DG.; Carrasco, M. Covert attention enhances letter identification without affecting channel tuning; Journal of Vision. 2004. p. 3p. 22-31.http://www.journalofvision.org/content/4/1/3

Todd JT, Norman JF, Mingolla E. Lightness constancy in the presence of specular highlights. Psychological Science. 2004; 15:33–39. [PubMed: 14717829]

Treue S. Perceptual enhancement of contrast by attention. Trends in Cognitive Sciences. 2004; 8:435–437. [PubMed: 15450502]

Tsal Y, Shalev L, Zakay D, Lubow RE. Attention reduces perceived brightness contrast. Quarterly Journal of Experimental Psychology A. 1994; 47:865–893.

Turatto M, Vescovi M, Valsecchi M. Attention makes moving objects be perceived to move faster. Vision Research. 2007; 47:166–178. [PubMed: 17116314]

Valsecchi M, Vescovi M, Turatto M. Are the effects of attention on speed judgments genuinely perceptual? Attention, Perception & Psychophysics. 2010; 72:637–650.

Van der Burg, E.; Olivers, CN.; Bronkhorst, AW.; Theeuwes, J. Audiovisual events capture attention: Evidence from temporal order judgments; Journal of Vision. 2008. p. 2p. 1-10.http://www.journalofvision.org/content/8/5/2

von Helmholtz, HLF. Handbuch der physiologischen Optik. Leipzig: Voss; 1910.

Wichmann FA, Hill NJ. The psychometric function: I. Fitting, sampling, and goodness of fit. Perception & Psychophysics. 2001a; 63:1293–1313. [PubMed: 11800458]

Wichmann FA, Hill NJ. The psychometric function: II. Bootstrap-based confidence intervals and sampling. Perception & Psychophysics. 2001b; 63:1314–1329. [PubMed: 11800459]

Williford T, Maunsell JH. Effects of spatial attention on contrast response functions in macaque area V4. Journal of Neurophysiology. 2006; 96:40–54. [PubMed: 16772516]

Yantis, S. Goal-directed and stimulus-driven determinants of attentional control. In: Monsell, S.; Driver, J., editors. Attention and performance. Vol. 18. Cambridge, MA: MIT Press; 2000. p. 73-103.

Yeshurun Y, Carrasco M. Attention improves or impairs visual performance by enhancing spatial resolution. Nature. 1998; 396:72–75. [PubMed: 9817201]

Yeshurun Y, Carrasco M. Spatial attention improves performance in spatial resolution tasks. Vision Research. 1999; 39:293–306. [PubMed: 10326137]

Appendix A

Ideal Observer Model

To better understand how our function parameters are influenced by different criteria, we implemented a signal detection theory model with an ideal observer. We simulated 10,000 runs of the comparative and the equality experiment (neutral condition) with 20 trials per data point. For the simulations, we mapped the contrast axis onto d'-values spanning from −4.65 to +4.65. These values correspond approximately to the theoretical maximal d'-values (assuming that in 100 trials observers make one false alarm). We assumed a Gaussian noise distribution centered at zero, corresponding to the standard stimulus contrast, and a Gaussian signal distribution for each test contrast centered at the corresponding d'-value. We assumed unit standard deviation for all noise and signal distributions and computed ideal criteria based on the true odds of signal and noise trials.

For the equality paradigm, we assumed one signal distribution for low and one for high contrasts with means equal to the average d' across test contrasts lower than and higher than standard contrast, respectively. For the comparative paradigm, the ideal criterion is exactly at zero. Then, the simulated data from each run were fit by the same cumulative/scaled Gaussian model that was used for the empirical data, and we derived the same set of parameters for each run. In a second simulation, we calculated criteria that were shifted according to the cube root of the real odds (Green & Swets, 1966; Maloney, 2002), which corresponds to assuming that signal ("different") trials occur with a probability of 0.6753. Note that this adjustment of criteria applies to the equality paradigm only, where the odds of noise and signal trials were not balanced.

Figure A1 shows the distributions of parameter estimates derived from the ideal observer model in the equality and the comparative experiments. In both experiments, the distribution of the PSE parameter is centered at zero, which corresponds to the true PSE. The standard

deviation parameter is slightly larger than 1 (mean = 1.13) in the ideal observer equality simulation, but is correctly estimated as ~1 (mean = 0.97) in the comparative simulation. Note that the amplitude parameter in the equality simulation converges on 0.65, meaning that an ideal observer knowing the odds of same and different trials only responds "same" in ~65% of the trials in which the stimulus contrasts are actually the same. Adjusting observers' criterion to the extreme probabilities of occurrence of "same" trials in the equality experiment (Green & Swets, 1966; Maloney, 2002) does not affect the PSE but results in larger standard deviations (mean = 1.22) and higher amplitudes (mean = 0.77), which match the empirical data in this study more closely than the ideal observer simulation. Note that whereas in the empirical data, the variability of the PSE estimates was greater for the equality than the comparative judgment, the variability in the simulations was similar for both judgments. This further indicates that the real observers do not behave ideally. Whereas part of the variance in the real data can be attributed to different criterion settings between observers, criterion settings were fixed across repetitions in each simulation.

In sum, this analysis allows us to test how different criteria settings influence parameter estimates in the equality paradigm. This result suggests that observers react to the extreme probabilities of "same" and "different" trials by developing a bias for responding "same."

Note

We also fit the data from the equality experiments with a similar model to the one Schneider and Komlos (2008) used, that is, the difference of two cumulative Gaussian functions of the form

$$0.5 * \mathrm{erf}\left(\frac{T - x - C}{SD\sqrt{2}}\right) - \mathrm{erf}\left(\frac{-T - x - C}{SD\sqrt{2}}\right), \tag{A1}$$

where T is the criterion, C is the contrast at the center of the function, and SD is the standard deviation of each Gaussian. This type of model is problematic for several reasons: It assumes symmetric criteria on both sides of the standard contrast and equal slopes on both flanks of the distribution. Both of these assumptions are unwarranted. However, if criteria are allowed to vary independently, the location (center) parameter becomes degenerate— basically the model would then assume that attention can only shift the distribution by affecting the criteria, which is a strong and unjustified theoretical assumption. Also, it allows for ill-defined, plateau-like peaks, which renders PSE estimation unreliable.

Nevertheless, we tested this model to directly compare our results with those of Schneider and Komlos (2008). This model on average did not give significantly different fits than the scaled Gaussian model (2 models × 7 increments ANOVA, no main effect of model, $F(1,8)$ < 1; 2 models × 3 cue conditions ANOVA, no main effect of model, $F(1,8) = 1.827$, $p = 0.213$), but it would require us to exclude two observers with at least one $R^2 < 0.7$ instead of the one observer we excluded from the increment experiment. Given that the scaled Gaussian model gave reasonably good fits and requires fewer theoretical assumptions, we used it for all further analyses.
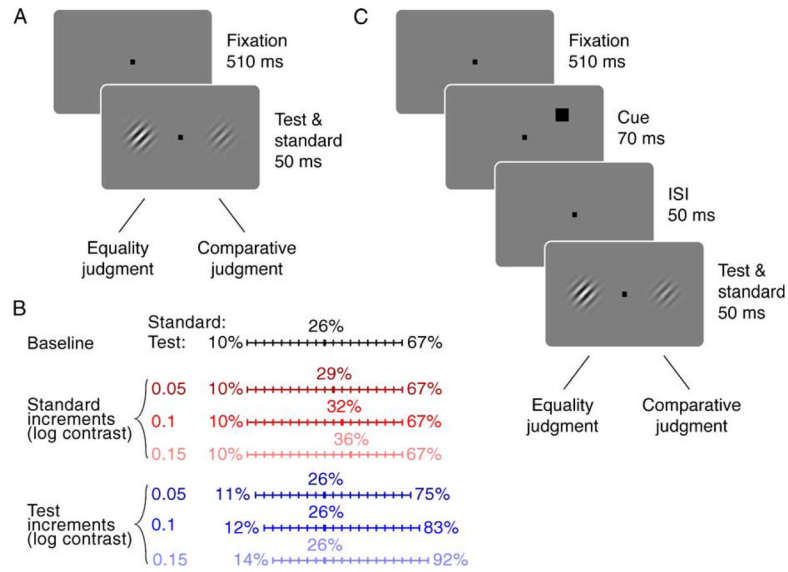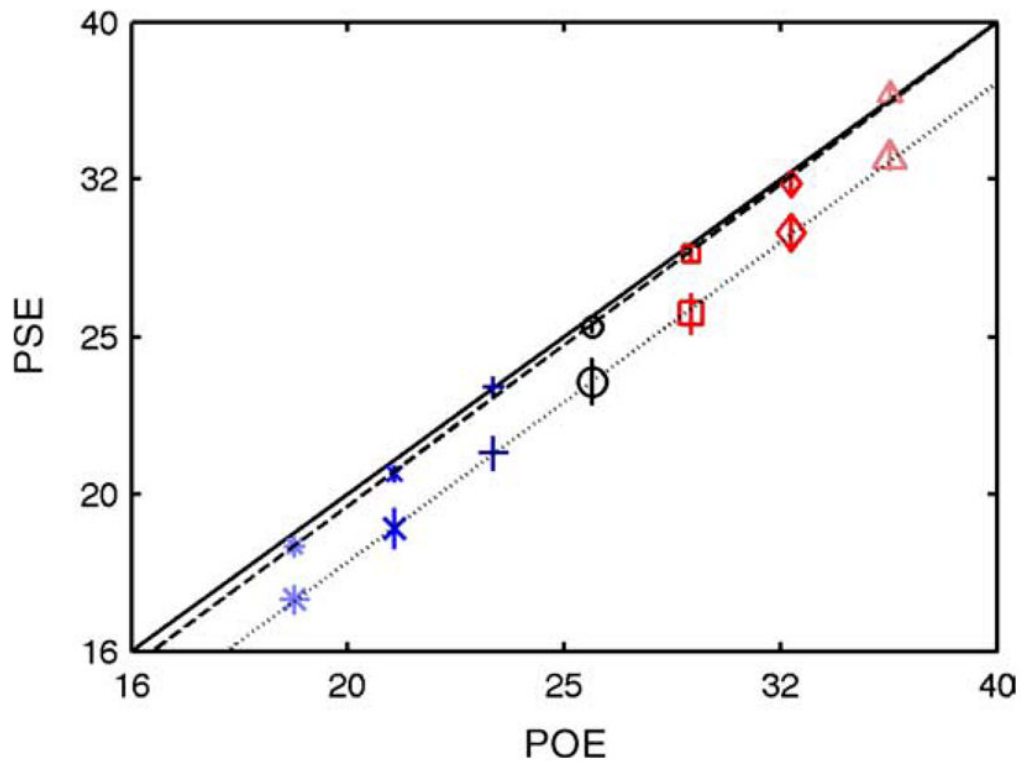
Figure 1.
Experimental design. (A) In both increment Experiments (1 and 2), each trial started with a fixation period of 510 ms; then two Gabors were presented for 50 ms at 4 deg eccentricity left and right of fixation. In the equality judgment (Experiment 1), observers were asked to indicate if the contrasts of the two stimuli were the same or different, while in the comparative judgment (Experiment 2), they had to report which stimulus had higher contrast. (B) Illustration of the increment conditions in Experiments 1 and 2. For each, the numbers on the left show the increment magnitude (log contrast), the axes on the right illustrate the resulting standard (thick tick mark) and test contrast (thin tick marks). No increment (baseline, black), test incremented by 0.05 (dark blue), 0.1 (medium blue), 0.15 (light blue) log contrast steps, or standard incremented by 0.05 (dark red), 0.1 (medium red), or 0.15 (light red) log contrast steps. (C) In the attention Experiments (3 and 4), the fixation period was followed by a cue which was flashed for 70 ms at fixation (neutral condition) or at 1.5 deg above the center of one of the Gabors (test cued and standard cued conditions). After an interstimulus interval of 50 ms, the two Gabors were presented for 50 ms, and observers reported if the contrasts were the same or different or which stimulus had higher contrast in the equality (Experiment 3) and comparative (Experiment 4) judgments, respectively. Drawings are not to scale.
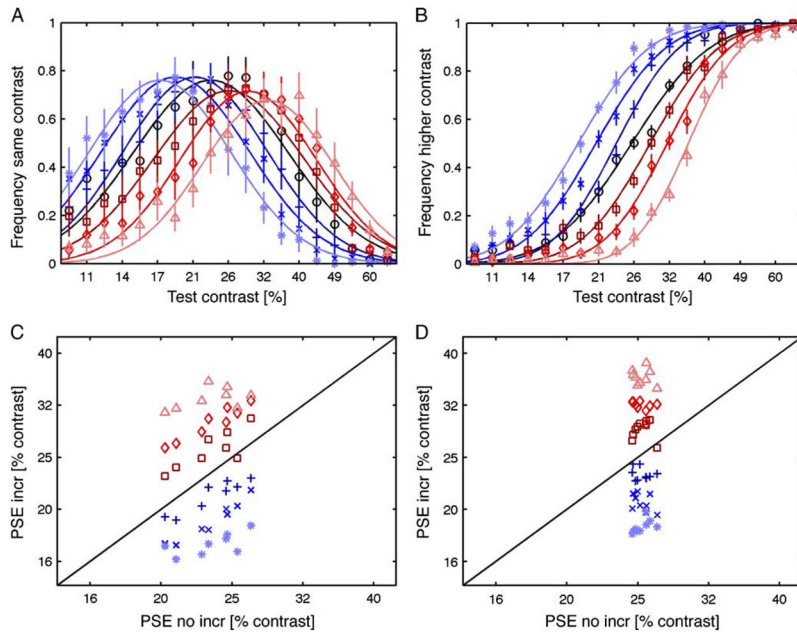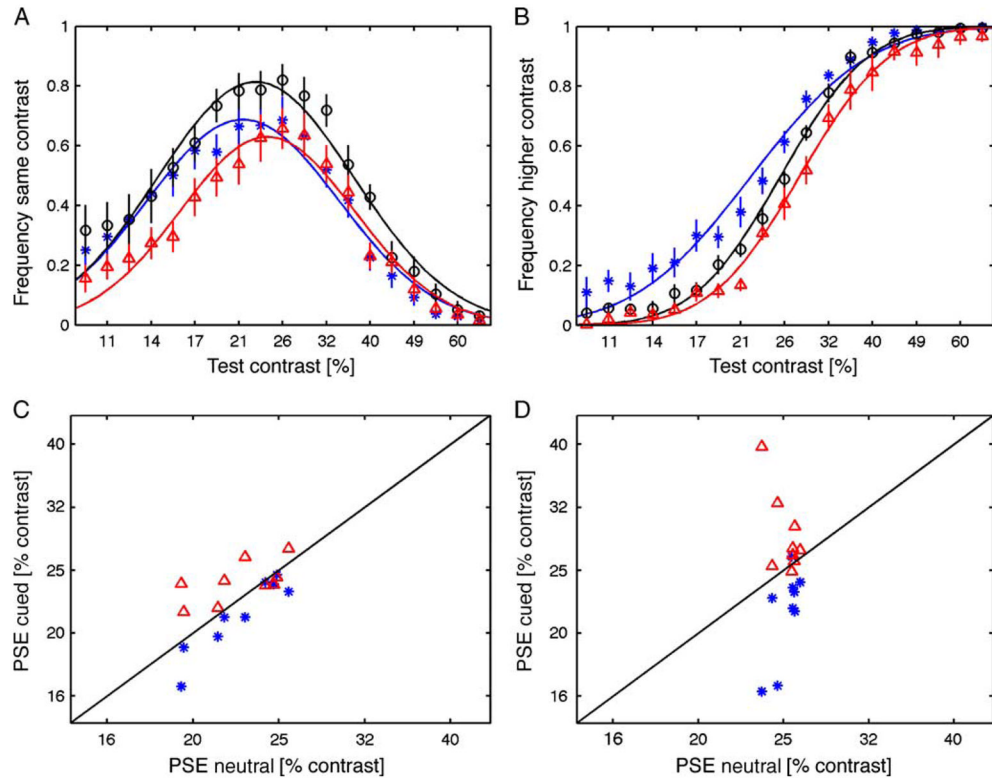
Figure 2.
Results Experiments 1 and 2 (Physical increments). (A) Experiment 1: Frequency of reporting same contrast as a function of no-increment test contrast (baseline) for each increment condition averaged across observers. Increment conditions: test incremented by 0.15 (light blue, *), 0.1 (medium blue, ×), or 0.05 (dark blue, +) log contrast steps, no increment (black, ○), standard incremented by 0.05 (dark red, □), 0.1 (medium red, ◆), or 0.15 (light red, △) log contrast steps. Error bars are standard error of the mean. Solid lines are fits to the average data. (B) Experiment 2: Frequency of reporting higher contrast as a function of baseline test contrast for each increment condition averaged across observers. Same format as A. (C and D) PSE in the test-incremented and standard-incremented conditions as a function of the PSE in the no increment condition for each observer in the equality (C) and comparative (D) experiment. Same symbols and color code as in A and B.
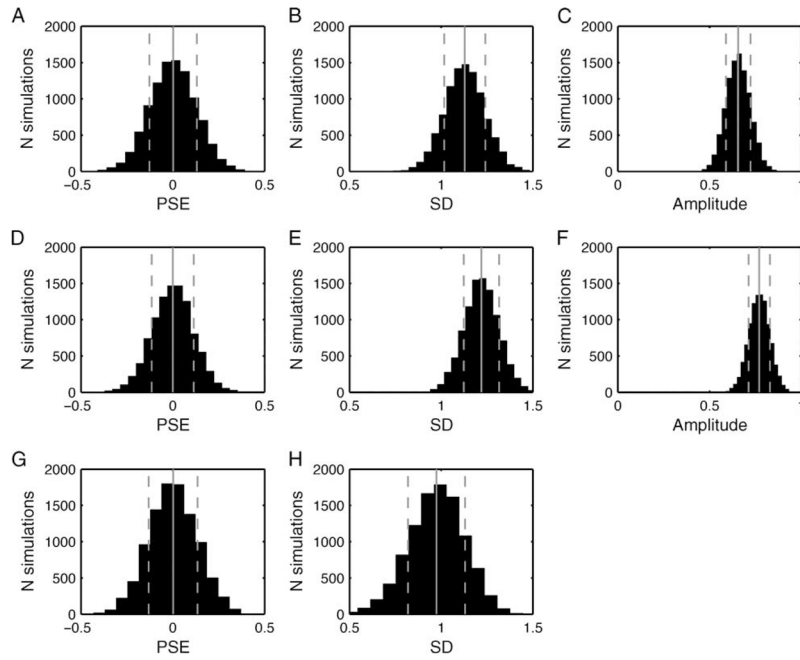
Figure 3.
PSE estimates Experiments 1 and 2 (Physical increments). Average PSE as a function of POE for each increment condition from the equality (large symbols) and comparative (small symbols) task. Lines mark the linear regression to the average data (dotted: equality, dashed: comparative experiment). Same symbols and color code as in Figure 2. Error bars are standard error of the mean.

Figure 4.
Results Experiments 3 and 4 (Attention). (A) Experiment 3: Frequency of reporting same contrast as a function of test contrast for test-cued (blue, *), neutral (black, ○), and standard-cued (red, △) conditions, averaged across observers. Error bars are standard error of the mean. Solid lines are fits to the average data. (B) Experiment 4: Frequency of reporting higher contrast as a function of test contrast for test cued (blue, *), neutral (black, ○), and standard-cued (red, △) conditions, averaged across observers. Same format as A. (C–D) PSE in the test-cued and standard-cued conditions as a function of the PSE in the neutral condition for each observer in the equality (C) and comparative (D) experiment. Same symbols and color code as in A and B.

Figure A1.
Simulation results. Top row: Equality experiment with ideal observer; distribution of PSE (A), standard deviation (B), and amplitude (C) estimates from 10,000 simulations with 20 trials per contrast level and ideal observer criterion. Medium row: Equality experiment with adjusted criterion; distribution of PSE (D), standard deviation (E), and amplitude (F). Bottom row: Comparative experiment; distribution of PSE (G), and standard deviation (H). Solid lines mark the mean, dashed lines one standard deviation. See text for details.

Table 1

Average parameter estimates with standard error across observers for the increment experiments. The POE is given in units of log contrast (leftmost column) and percent contrast (in parentheses). The remaining cells reflect the mean (PSE), standard deviation (*SD*), and amplitude parameters for the equality paradigm, and then the mean (PSE) and standard deviation (*SD*) parameters for the comparative paradigm. In each cell, the top row refers to the parameter estimate in units of log contrast and the bottom row to the standard error (in italics); for the PSE, the corresponding contrast value in percent contrast is given in parentheses.

| Parameter | POE | Equality (Experiment 1) *n* = 8 | | | Comparative (Experiment 2) *n* = 9 | |
		PSE	SD	Amplitude	PSE	SD
Test increment 3	−0.725 (19)	−0.767 (17.1)	0.173	0.776	−0.733 (18.5)	0.147
		0.021	*0.048*	*0.083*	*0.013*	*0.052*
Test increment 2	−0.679 (21)	−0.722 (19.0)	0.181	0.788	−0.687 (20.6)	0.145
		0.037	*0.045*	*0.082*	*0.017*	*0.043*
Test increment 1	−0.633 (23)	−0.674 (21.2)	0.180	0.780	−0.632 (23.3)	0.134
		0.032	*0.046*	*0.077*	*0.012*	*0.044*
No increment	−0.587 (26)	−0.629 (23.6)	0.188	0.778	−0.594 (25.5)	0.162
		0.042	*0.053*	*0.083*	*0.012*	*0.044*
Standard increment 1	−0.541 (29)	−0.586 (26.0)	0.186	0.732	−0.547 (28.4)	0.146
		0.036	*0.039*	*0.089*	*0.018*	*0.061*
Standard increment 2	−0.495 (32)	−0.534 (29.3)	0.164	0.742	−0.503 (31.4)	0.133
		0.032	*0.049*	*0.086*	*0.014*	*0.053*
Standard increment 3	−0.449 (36)	−0.487 (32.6)	0.162	0.689	−0.446 (35.8)	0.115
		0.021	*0.044*	*0.092*	*0.016*	*0.046*

Table 2

Average parameter estimates with standard error across observers for the attention experiments. The cells reflect the mean (PSE), standard deviation (*SD*), and amplitude parameters for the equality paradigm, and then the mean (PSE) and standard deviation (*SD*) parameters for the comparative paradigm. In each cell, the top row refers to the parameter estimate in units of log contrast and the bottom row to the standard error (in italics); for the PSE, the corresponding contrast value in percent contrast is given in parentheses.

Parameter	Equality (Experiment 3) $n = 9$			Comparative (Experiment 4) $n = 9$	
	PSE	SD	Amplitude	PSE	SD
Test cued	−0.672 (21.3)	0.198	0.714	−0.667 (21.5)	0.193
	0.058	*0.036*	*0.068*	*0.073*	*0.075*
Neutral cue	−0.646 (22.6)	0.206	0.828	−0.596 (25.3)	0.147
	0.047	*0.050*	*0.055*	*0.015*	*0.056*
Standard cued	−0.619 (24.1)	0.186	0.641	−0.546 (28.4)	0.145
	0.032	*0.036*	*0.074*	*0.063*	*0.051*