Confidence as a diagnostic tool for perceptual aftereffects

Regan M. Gallagher[1,2], Thomas Suddendorf[1], & Derek H. Arnold[1]

[1]School of Psychology, The University of Queensland,

St Lucia, Brisbane, Australia.

[2]Cognitieve Psychologie, Vrije Universiteit Amsterdam

Amsterdam, The Netherlands.

Total word count: 5,508

Abstract: 143
Intro & Discussion: 1,932
Figures & Tables: 740 [6 Figs, 1 Table]
Methods & Results: 1,796
References: 492 [22 refs]

Corresponding author: Regan Gallagher
Email: regan.gallagher@uqconnect.edu.au

Abstract

Much of what we know about how our sensory systems operate comes from studying aftereffects. Prolonged and repeated exposure to a constant stimulus (adaptation) can impact responses to subsequent test stimuli, compared to a non-adapted test stimulus. Researchers often infer, based on changes to response patterns, that sensory encoding has been changed by adaptation. However, it can be ambiguous whether a given aftereffect changes sensory encoding, decision processes, or both. We propose that confidence estimates can be used as a diagnostic tool to identify whether aftereffects arise either partially or entirely from post-perceptual processes. In three experiments, we demonstrate the utility of confidence judgements in determining whether an aftereffect results from sensory or criterion changes. Given the simplicity of our approach, this tool could readily be applied to resolve contemporary controversies regarding the origins of aftereffects and the penetration of perception by cognition.

Key words: Confidence, Adaptation, Perceptual Aftereffects

A central challenge in perception research is to understand how the world looks, feels, and sounds, as opposed to how it is remembered, imagined, or judged. Much of what we know about human perception is the result of investigating sensory aftereffects (e.g. Mollon, 1974; Wark, Lundstrom, & Fairhall, 2007). Aftereffects are usually quantified by recording subjective responses to a range of stimulus intensities (e.g. brightness of a light, frequency of a tone), and measuring decision changes before and after prolonged and repeated exposure (see Clifford, Webster, Stanley, Stocker, Kohn, & Sharpee et al., 2007 for a review). Responses collected using psychophysical methods are, by nature, an alloy of sensory functions and the cognitive systems responsible for producing measurable behaviour. This can create ambiguity regarding the cause of response changes.

Signal detection theory (SDT; Tanner & Swets, 1954) and forced-choice methods can be used to measure objective sensitivity to physical stimulations. They are also effective at minimising the influence of biases arising from cognitive processes. But aftereffects are often subjective, and may or may not be accompanied by changes in sensitivity (for greater discussion on this topic, see Morgan et al, 2011; Storrs, 2015). In these cases, it can be difficult for SDT to parse a decision change into its sensory and cognitive components. Consequently, it is often difficult to determine at which stage of processing an aftereffect arises and whether a measured change in behaviour is due to a change in perception, cognition, or both.

When measuring subjective aftereffects, adapting and test stimuli are typically presented in the same sensory domain, such as a moving adaptor influencing the perceived motion of a test (Barlow & Hill, 1963). However, recent research has started revealing aftereffects that are only conceptually related to the adapting stimulus. For example, in one study conducted by Winawer, Huk & Boroditsky (2008), participants adapted to still photographs that implied motion

either leftward or rightward, or inward or outward. In a second study (Winawer, Huk & Boroditsky, 2010), participants adapted to a static grating stimulus and imagined it to be moving either upward or downward. Participants then judged the motion direction of a coherent dot stimulus. In both studies, the adaptation phase of the task gave rise to a negative aftereffect (participants were more likely to judge an ambiguous stimulus as having moved in the direction opposite to the implied or imagined direction), consistent with results produced by the classic motion aftereffect. The authors then concluded that the adaptation task had impacted motion perception. While this interpretation is intuitive, there is, however, another plausible interpretation. Viewing implied motion, imagining real motion, or activating other motion-related cognitions, might instead impact *decisions* about the direction of moving tests, yet leave the processes underlying motion perception unchanged.

Firestone and Scholl (2013) investigated a similar claim by Banerjee et al. (2012) that moral judgements affect brightness perception. In one experiment, they directly replicated the original study in which participants wrote about a moral or immoral action from their past, and then judged the brightness of the room they were in on a 7-point scale. Participants in the immoral condition rated the room's brightness as lower than participants in the moral condition. This result is broadly consistent with a positive aftereffect (with thinking about 'dark' moral behaviours inducing a perceived darkening of the environment). A potential problem with this interpretation is the El Greco fallacy, named after the 16th century artist known for painting strikingly elongated figures in his portraits. Some claimed this tendency was caused by astigmatism, which results in unfocused and elongated retinal projections. This intuitive explanation is a fallacy; any such visual distortion would apply, in equal proportion, to viewed objects and the canvas on which they are painted (for a conceptual primer on why this is true, see Anstis,

2002; for a discussion of constant-error distortions in psychophysics, see Morgan et al., 2014).

In keeping with this reasoning, Firestone & Scholl (2013) conducted a follow-up study, replacing the 7-point scale with seven grey-scale images with linearly decreasing brightness. The authors reasoned that if the world looked darker after dwelling on immoral actions, the effect would apply equally to the room as well as the brightness of the grey scale, thereby cancelling out any effect. However, participants in the immoral condition again rated their room as being darker (i.e. matching their room to a less bright image on the grey scale) compared with participants in the moral condition. This finding suggests that Banerjee and colleagues were not measuring changes to participants' lightness perception, so much as measuring a decision change after dwelling on moral quandaries (to think otherwise is to commit the El Greco fallacy, see Firestone & Scholl, 2015). This example highlights the distinction between changed decisions and changed perceptions.

Perceiving versus deciding

Figure 1 depicts two components of perceptual decisions in a binary categorisation task — sensory evidence and decision criteria (Tanner & Swets, 1954). Sensory evidence consists of the physiological information available to the brain from the sense organs, and these signals are contaminated by sensory noise. Criteria can be regarded as decision boundaries, in this case for categorising inputs as moving to the left or to the right. Perception of weaker stimuli (with encoded values close to a criterion) can be shifted to either side of the criterion on a trial-by-trial basis by sensory noise, so decisions made about weak signals are probabilistic. If the observer is forced to decide between two alternatives (e.g. motion to the left or right) when an input is perfectly ambiguous, they could be forced to rely on higher-order cognitive (i.e. post-perceptual) aspects of decision-

making. An observer might simply guess, leading to an equal likelihood of choosing left or right. Alternatively, they may adopt some other strategy that carries greater bias toward one category or the other. A decision strategy could manifest as either a negative or a positive bias (or none), independent of sensory evidence and other post-perceptual processes (see Yarrow et al, 2011), and changes to decision boundaries need not change objective sensitivity (Morgan et al., 2011).

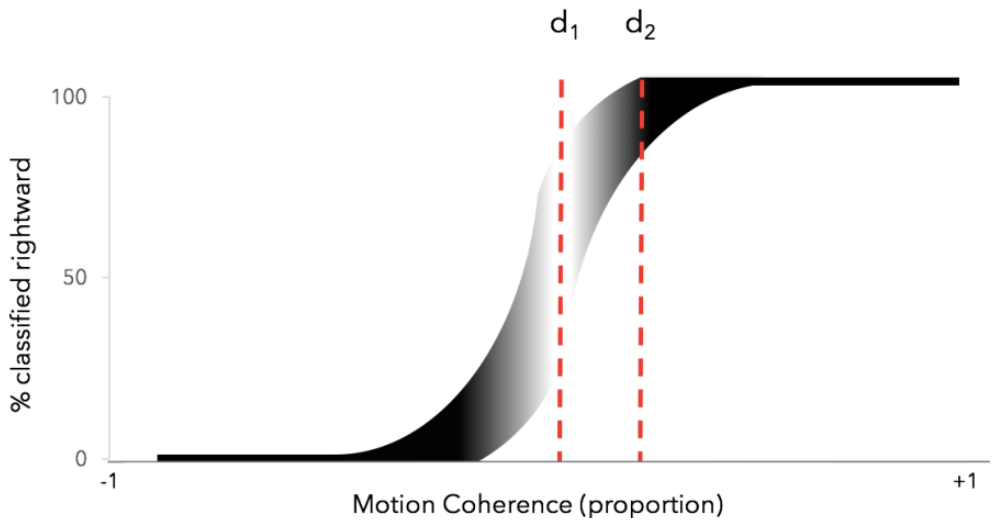

Figure 1: A sigmoid function (S-curve) representing the sensory signal-to-noise ratio (greater noise implied by lighter shading) and decision criteria (red dotted lines) in motion direction discriminations across a continuum of stimulus intensities. As the stimulus intensity approaches 0, the signal-to-noise ratio decreases and the observer's decisions become more probabilistic. The placement of the criterion determines the point of subjective equality (PSE), or the stimulus magnitude where the sensory signal is judged leftward and rightward with equal probability. A criterion placed at d_1 indicates equal likelihood for choosing either response when sensory evidence is ambiguous. A criterion placed at d_2 indicates a tendency toward choosing 'leftward' in the absence of compelling sensory evidence for the 'rightward' category. Placement of a criterion therefore directly impacts PSE estimates independent of sensory evidence.

Confidence as a diagnostic tool

If an aftereffect changes the way the world looks, we would regard it as a perceptual aftereffect. The motion aftereffect, for example, can make still landscapes or non-coherent motion appear to be moving. This is best explained

by physiological changes to the processing of motion energy (Barlow & Hill, 1963; Mollon, 1974). An observer, after adaptation, will confidently (but incorrectly) report a static stimulus is moving. By corollary, the point at which the observer is forced to guess the motion direction will also change — a stimulus moving slowly in the direction of the adapting stimulus would now appear stationary — resulting in probabilistic forced-choice responses. This pattern of behaviour is depicted by the 'Encoding component' of sensory adaptation (see Figure 2). The same pattern of decision changes could occur without physiological changes, not because the world looks different, but due to a shift in decision criteria. A participant whose task is to imagine or infer movement while viewing stationary objects might, when prompted, show a statistical preference for the unimagined direction when the test stimulus is ambiguous (Winawer et al., 2010, but see Morgan et al., 2011). In this example, the observer is judging the same physiological evidence with a different criterion. This is the 'Criterion component' of an aftereffect (also in Figure 2).

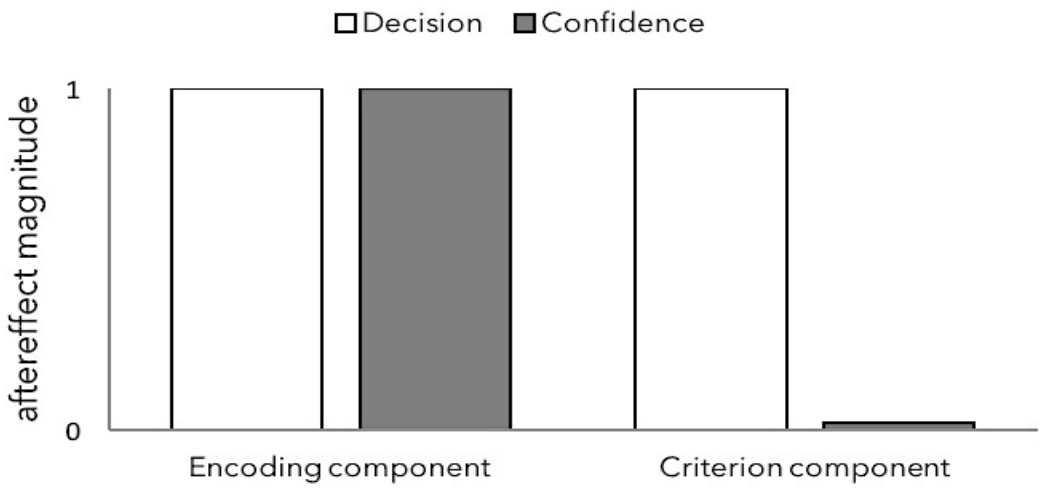

Figure 2: The magnitude of an aftereffect is often measured by the change in the point of subjective equality (ΔPSE) following adaptation. Here these are represented by a value of 1. On the basis of a PSE shift alone, it is not clear whether an aftereffect arises from changes to sensory processing or from a criterion shift. The Encoding component implies decision changes reflect changes to sensory evidence. In this case, a decision change is associated with a commensurate change in an observer's uncertainty. The criterion component implies decision changes result from a tendency to favour a particular decision when presented with ambiguous sensory evidence.

Therefore, we predict that adaptation would create a robust change in direction decisions for particular signal intensities, but these decision changes would occur without changes in the range of stimuli that elicit perceptual uncertainty, as measured by confidence ratings.

Confidence judgements might offer insight into the role of decision criteria in perceptual aftereffects. Aftereffects caused by physiological changes should equally impact decision and confidence estimates because the sensory evidence underlying both judgements has changed. Aftereffects arising from criterion changes, however, might selectively impact decisions made under uncertainty. In this case, observers should report no increase in confidence for physically ambiguous stimuli, nor should they become less confident about decisions for which they were previously confident.

Aim and hypotheses

The aim of the current study is to test whether estimates of confidence can provide diagnostic information about the cause of an aftereffect. To do this, we used three methods of adaptation and then measured motion-direction decisions. We also measured participants' confidence in their direction judgements. We hypothesise that perceptual aftereffects will impact both decision and confidence reports to the same degree. In contrast, we hypothesise that aftereffects arising from post-perceptual processes will bias direction decisions, but not reported confidence.

Method

Participants

All participants (15, 30, and 15 for Experiments 1–3, respectively) were recruited from the University of Queensland's Psychology department and each person provided verbal consent. Sample sizes of 15 were set for Experiments 1 and 3, as sensory aftereffects can be regarded as large (Cohen's d of >= 0.8), so these sample sizes should ensure an above 80% chance of successfully measuring an effect. A more conservative sample size of 30 was chosen for Experiment 2, as this may examine a weaker medium-sized cognitive effect. This sample size should ensure an ~80% chance of successfully measuring a previously reported effect (see Winawer, Huk & Boroditsky, 2008) with a comparable sample size.

In Experiments 1 and 2, participants were drawn from a first-year student pool and received course credit for their time. Participants in Experiment 3 were postgraduate Psychology and UQ Perception Lab members. All were naïve to the purpose of the experiments. Ethical approval for all experiments was obtained from the University of Queensland's Ethics Committee, and were in accordance with the Declaration of Helsinki.

Materials

Stimuli in Experiments 1 and 2 were presented on an LCD monitor running Windows XP. Experiment 3 stimuli were presented on a CRT monitor running Mac OSX. All computers were running Matlab software and the Psychophysics Toolbox (Brainard, 1997; Pelli, 1997). All monitors had a screen refresh rate of 60Hz.

Stimuli

Test stimuli consisted of 100 dots rendered blue against a grey background. Each dot was 1 pixel in size. A subset (equal to the coherence of the test stimulus) of dots was re-drawn one pixel from their previous location, in the

coherent direction, on successive frames. Other dots were redrawn at random locations. The coherently moving dots were selected at random on each frame, so no individual dot could be tracked across the screen. Dot coherence values ranged from -30 (30% coherence leftward) through 0 (random motion) to +30 (30% coherence rightward). Test stimuli were set to one of 11 coherence values [-30 -20 -10 -6 -3 0 3 6 10 20 30], presented in a randomised order. All stimuli were drawn within a window 33% the height and width of the monitor.

In Experiment 1, adapting stimuli consisted of 100 dots rendered red, each 1 pixel in size. Coherent motion was achieved by selecting 30 dots at random on each frame (to prevent tracking individual dots) to be displaced left or right by one pixel. All other dots were redrawn at random locations. Adapting coherent motion was in one direction for the first half of the block of trials, and in the other direction for the second half of the block.

Adapting stimuli in Experiment 2 were photographs that implied motion, either to the left or to the right. A collection of 100 photographs comprised the set of implied motion adaptors, chosen at random on each trial. Each photo was mirror-flipped, so as to depict both directions. Experiment 3 adapting stimuli were the same as for Experiment 1, with two exceptions. First, motion coherence was always set to 0, and second, a static direction cue was presented directly above the stimulus window — an arrow pointing to the left or right (see Figure 3). The static direction cue was consistent within each block of trials.

Procedure

Participants sat comfortably in a chair approximately 55 cm away from the display, rested their hands on the keyboard's directional buttons, and fixated a central cross-hair. The adapting stimulus was presented for 18 seconds on the first trial for each of five blocks, and again on the middle trial when the motion direction reversed. For all other trials the adapting stimulus lasted for six seconds.

Participants passively viewed the coherent motion (Exp. 1) implied motion (Exp. 2) or random motion with visual cue (Exp. 3; see Figure 3 for a representation of the task procedure).

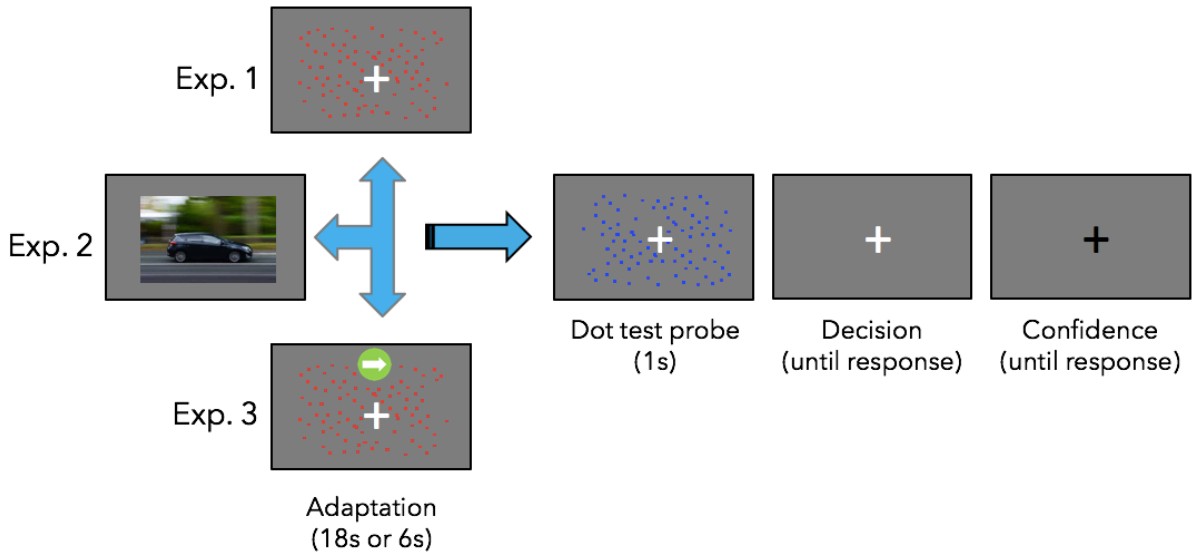

Figure 3: Experimental procedures. For each experiment, participants viewed stimuli that depicted a) coherent motion, b) implied motion, or c) random motion and a static direction cue. Each trial within a block consisted of an adaptation phase followed by a dot test probe. The direction of the adapting stimulus (left or right) was consistent within the first half of a block, and then changed direction for the second half. Adapting stimuli appeared for 18s on the first trial of each block, and on the middle trial when the stimulus changed direction, and for 6s on all other trials. Dot test probes were present for 1s, appearing on the second frame after the adapting stimulus disappeared. A new trial began once participants had recorded their direction decision (left or right) and reported their confidence (whether they believed their decision was correct — yes or no).

After the adaptation phase, participants were presented with one of the 11 dot-motion test probes. The test was visible for one second before disappearing, leaving only the fixation cue. Participants recorded whether the test had appeared to be moving left (by pressing the left arrow key) or right (by pressing the right arrow key). In Experiments 1 and 2, participants were instructed to provide their

best guess if they could not determine which direction the test was moving. If participants could not determine test direction in Experiment 3, they were instructed to default their response to the direction congruent (Experiment 3a) or incongruent (Experiment 3b) with the static cue. Participants in Experiment 3 completed both tasks (congruent and incongruent) in a randomised and counterbalanced order.

Once the direction judgment had been made on each trial, the fixation cue turned black, prompting a confidence response. Participants indicated whether they felt their response was correct by pressing the up arrow on the keyboard (high confidence response) or probably wrong by pressing the down arrow (indicating low confidence in their decision, or that they were guessing). The fixation cross turned white once the confidence response had been provided, and a new trial started immediately. Each of the 11 stimulus values was tested twice per block (once with a leftward adaptor and once with a rightward adaptor). The adaptor preceded each of the 11 values in one direction, and then each of the 11 values in the other direction. The initial adaptor direction was randomised for each block of trials, and 5 blocks were completed by each participant during a single experimental session, resulting in 110 individual test trials being completed by each participant for each Experiment.

Results

In all three experiments, data for each participant produced two decision distributions, of apparent rightward motion as a function of test dot coherence and direction. One distribution describes responses for left-adapted trials, the other for right-adapted trials. Cumulative Gaussian functions were fit to each distribution, and 50% points were taken as estimates of the point of subjective equality (PSE) — the stimulus value equally likely to be judged as moving left or right. The impact of adaptation was estimated as the difference between the two PSE estimates for each individual. Data for each participant also produced two distributions of uncertainty (low-confidence) as a function of test dot coherence and direction, one for left-adapted trials, and one for right-adapted trials. A raised Gaussian function was fit to each distribution, and the peak of the fitted function was taken as an estimate of the point of maximal uncertainty. All *t*-tests reported are two-tailed repeated measures tests for equality of means.

Experiment 1

Coherent motion adaptation

Adaptation had a robust impact on direction decisions, with left adaptation (L_{PSE} = -11.67, SD = 7.31) and right adaptation (R_{PSE} = 11.22, SD = 6.76) producing different PSE estimates, *paired-t*(14) = 6.90, $p < .001$. The same was true for confidence data. Adaptation robustly impacted measures of uncertainty, with left adaptation (L_{CONF} = -10.02, SD = 5.79) and right adaptation (R_{CONF} = 9.93, SD = 5.27) producing different confidence profiles, *paired-t*(14) = 7.98, $p < .001$. Also, the size of the effect of adaptation on decision data (ΔPSE = 22.89, SD = 12.84) and confidence data (ΔCONF = 19.95, SD = 9.69) was statistically equivalent, *paired-t*(14) = 1.73, $p = .106$. Related data, averaged across participants, are depicted in Figure 4 (top row). This pattern of results is consistent with sensory adaptation affecting the Encoding component of an aftereffect

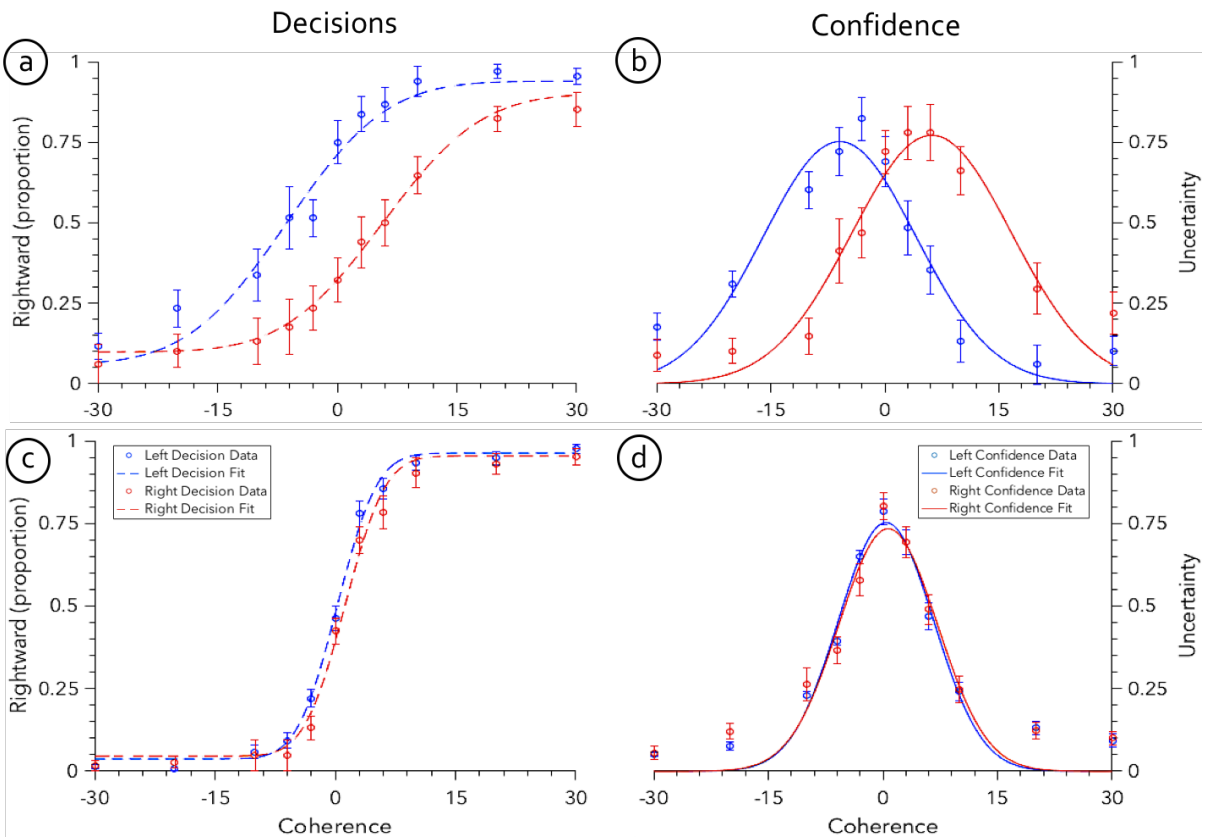

Figure 4: Perceived motion direction (left or right) and confidence (high or low) as a function of dot coherence (percent) and direction (negative values left, positive right). Data and best-fit functions are depicted following left (blue) and right (red) adaptation, for both coherent motion (Experiment 1, panels a and b) and implied motion (Experiment 2, panels c and d). *Left*: Decision data and function fits indicating PSE shifts, measured as the difference between the two decision curves. *Right*: Confidence data and function fits, indicating confidence shifts, measured as the difference between the peaks of uncertainty curves. All data are averaged across participants. Note, effects suggested by these data and function fits might differ subtly from results reported, as the latter were based on analyses of individual data sets. Error bars indicate ±1 S.E.

Experiment 2

Implied motion adaptation

Analyses for Experiment 2 were identical to those for Experiment 1. These showed that adaptation to still images depicting motion had a robust impact on direction decisions, with adaptation to left (L_{PSE} = 0.71, SD = 2.70) and right-implied motion images (R_{PSE} = 2.50, SD = 4.70) producing different PSE estimates, $t(29) = 2.94$, $p = .006$. However, implied motion adaptation did not impact on

measures of confidence, with left (L_{CONF} = 0.84, SD = 1.50) and right-implied (R_{CONF} = 0.78, SD = 2.29) motion images producing statistically equivalent estimates of peak uncertainty, $t(29) = 0.13$, $p = .896$. Indeed, the degree to which decision data were affected by implied motion (ΔPSE = 1.79, SD = 3.33) was robustly different to the degree that confidence estimates (ΔCONF = -0.06, SD = 2.55) were affected, $t(29) = 2.50$, $p = .018$. These data are depicted in Figure 4 (bottom row).

Given the conceptual importance of the confidence null result, we further assessed this using a Bayes factor analysis. This resulted in a Bayes factor of 0.02, assuming minimal and maximal effect sizes +/- 50% from the decision effect size. This provides strong evidence in favour of the null hypothesis, that there is no difference in points of maximal uncertainty between the two 'adaptation' conditions. Overall, our results are consistent with changes to the Criterion component of an aftereffect

Experiment 3

Random motion adaptation

Analyses for Experiment 3 were identical to those for Experiments 1 and 2. Experiments 1 and 2 established that aftereffects can have different metacognitive profiles. Both aftereffects, however, have previously been interpreted in terms of perceptual changes, rather than criterion changes. In Experiment 3 we eliminate the possibility of a perceptual direction aftereffect. To do this, participants adapted to a randomly moving stimulus, but were instructed to use a default response, either congruent (Experiment 3a) or incongruent (Experiment 3b) with a static direction cue if they were unable to determine test direction.

Cue-congruent task

In the cue-congruent task, participants defaulted their response to the direction of the cue if they could not determine the direction of the test stimulus. This pattern of response is expected to produce a bias in PSE data consistent with a positive aftereffect. Indeed, participants were more likely to say the test stimulus was moving left when the cue pointed left (L_{PSE} = 5.96, SD = 3.14), and more likely to say the stimulus was moving to the right when the cue pointed right (R_{PSE} = -7.50, SD = 3.48), $t(14) = 8.40$, $p < .001$, consistent with a positive aftereffect. These cued decisions, however, showed no sign of a significant impact on confidence data. Confidence estimates were neither biased substantially to the left when the cue pointed left (L_{CONF} = -0.12, SD = 1.13), or to the right when the cue pointed right (R_{CONF} = 0.62, SD = 1.07). Differences between left- and right-cued confidence estimates were non-significant, $t(14) = 1.15$, $p = .268$. The impact of the cue was greater for decision data (ΔPSE = 13.46, SD = 6.20) compared to confidence estimates (ΔCONF = 0.50, SD = 1.68), $t(14) = 9.23$, $p < .001$. These data are depicted in Figure 5 (top row).

Again, given the conceptual importance of the confidence null result, we conducted a Bayes factor analysis resulting in a factor of 0.01 (assuming minimal and maximal effect sizes +/- 50% from the average decision effect size). This provides strong evidence in favour of the null hypothesis, that there is no difference in points of maximal uncertainty between these two 'adaptation' conditions.

Cue-incongruent task

In the cue-incongruent task, participants were instructed to default to the direction opposite the cue if they could not determine the test direction. This was expected to produce PSE changes consistent with a negative aftereffect. Results were consistent with this prediction. Participants tended to respond that the

stimulus was moving to the right when the cue pointed to the left (L_{PSE} = -7.60, SD = 3.75), and respond that the stimulus was moving to the left when the cue pointed to the right (R_{PSE} = 7.33, SD = 4.65), resulting in different PSE estimates for these conditions, $t(14) = 7.31$, $p < .001$. Cue instructions once again, however, had no discernible impact on confidence data. Confidence estimates for left (L_{CONF} = 0.02, SD = 1.26) and right (R_{CONF} = -0.38, SD = 1.02) cues were statistically equivalent, $t(14) = 0.85$, $p = .410$. The impact of direction cues were greater for decision data (ΔPSE = 14.94, SD = 7.91) compared to confidence data (ΔCONF = 0.41, SD = 1.85), $t(14) = 7.07$, $p < .001$. These data are depicted in Figure 5 (bottom row). A Bayes factor analysis to assess the Confidence null result, resulting in a factor of <0.01 (assuming minimal and maximal effect sizes +/- 50% from the average decision effect size). This provides strong evidence in favour of the null hypothesis, that there is no difference in points of maximal uncertainty between these two 'adaptation' conditions.

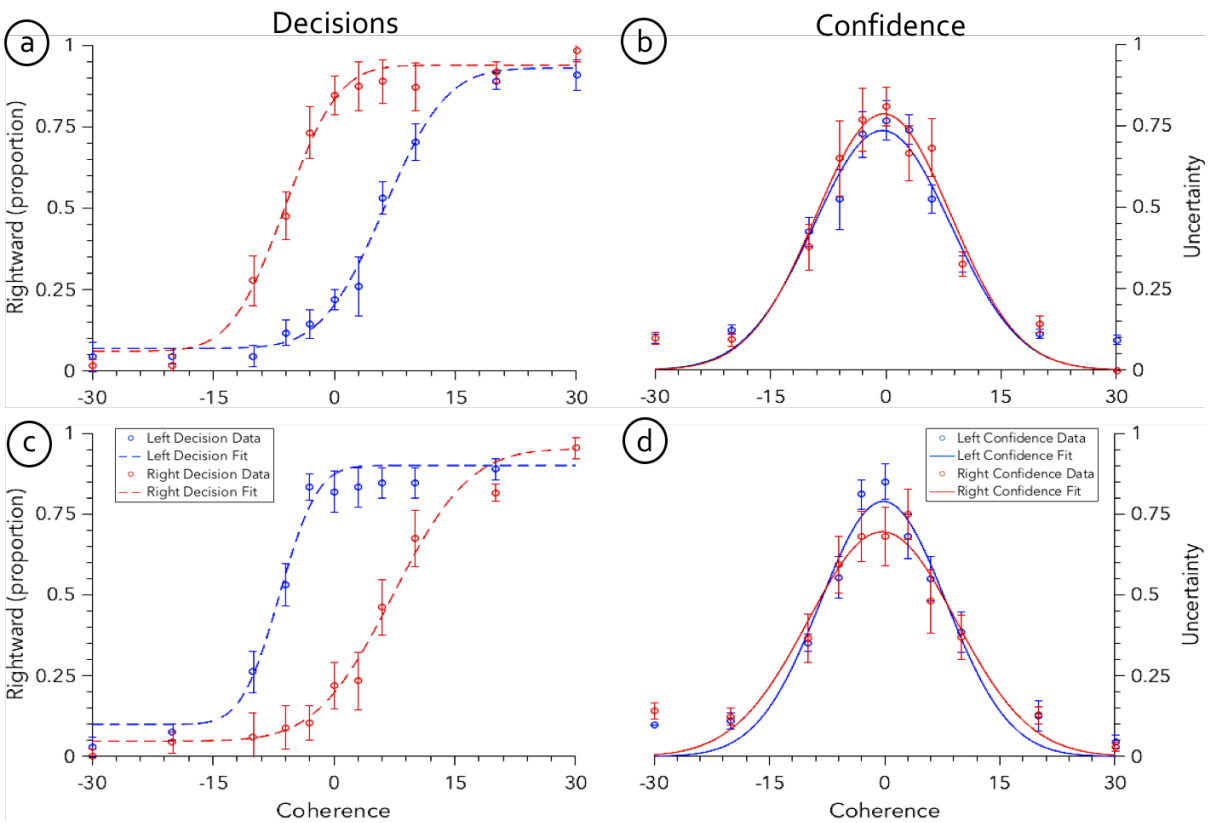

Figure 5: Perceived motion direction (left or right) and confidence (high or low) as a function of dot coherence (percent) and direction (negative values left, positive right). Data and best-fit functions are depicted following left and right criterion cues, for congruent (Experiment 3a, panels a and b) and incongruent (Experiment 3b, panels c and d) instructed biases. *Left*: Decision data and function fits indicating PSE shifts, measured as the difference between PSE estimates for the two decision curves. *Right*: Confidence data and function fits, indicating the confidence shift, measured as the difference between the peaks of uncertainty curves. All data are averaged across participants. Note, effects suggested by these data and function fits might differ subtly from results reported, as the latter were based on analyses of individual data sets. Error bars indicate ±1 S.E.

Normalised decision and confidence changes

Aftereffect magnitudes varied across experimental tasks. To aid comparison of results, aftereffect magnitudes were normalised by dividing measured individual aftereffects by the group average decision aftereffect (i.e. the magnitude of decision changes) in that experiment. Therefore, all normalised decision aftereffects have a mean of +1, and normalised confidence aftereffects are proportional to the decision aftereffects. As can be seen in Figure 6,

adaptation to coherent motion (Experiment 1) had equivalent effects on direction decisions and confidence reports, whereas adaptation to implied motion (Experiment 2) and instructing people to adopt a bias when uncertain about test direction (averaged for Experiment 3a and 3b) had unequal effects, with confidence proportions statistically equivalent to 0.

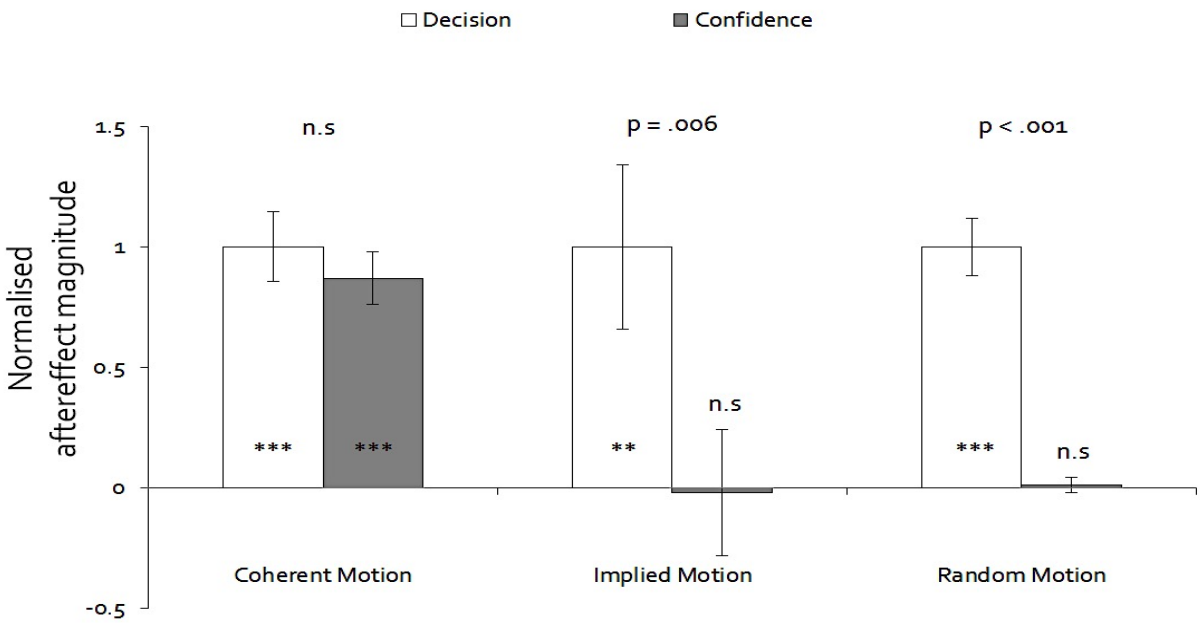

Figure 6: Decision and confidence aftereffects for each experiment, normalised to the average decision aftereffect size in each experiment. Error bars indicate ±1 S.E.

Sensitivity analyses

One possibility is that decision and confidence estimates were not equally sensitive. If this were true, a small aftereffect might be revealed by one measure, but not the other. To rule this out as a potential explanation for our findings, in each experiment we compared the precision of the two measures. We tested whether the distance between the 25 and 75% points of cumulative Gaussian functions (fit to decision data) was different to the full width at the half height of raised Gaussian functions (fit to confidence data). Note that these analyses relate to matched points on these two types of function fits. Analyses reveal that our

decision and confidence measures were equally precise (all *p*-values > .100; see Table 1 for means and standard deviations).

Table 1. Precision estimates (JNDs) for decision and confidence reports in each experiment.

	Left		Right	
	Decision	Confidence	Decision	Confidence
Exp1	5.78 (4.10)	5.15 (2.21)	7.02 (3.09)	6.71 (5.80)
Exp2	4.16 (3.00)	5.37 (4.93)	4.82 (3.36)	4.75 (3.18)
Exp3(a)	5.68 (3.44)	5.69 (3.71)	5.06 (2.40)	4.90 (2.21)
Exp3(b)	5.83 (5.39)	4.70 (3.09)	6.65 (3.99)	5.12 (2.24)

Mean coherence required to notice a difference, parentheses denote standard deviations.

General Discussion

Our data suggest that perceptual aftereffects have a metacognitive profile distinct from post-perceptual aftereffects. The classic motion aftereffect displayed an underlying symmetry between perceptual decision changes and confidence reports (Experiment 1). By contrast, implied motion adaptation (Experiment 2) changed direction decisions, but had no impact on confidence reports. Instructing participants to adopt a conscious decision strategy when uncertain (Experiment 3) also changed direction decisions, but not confidence reports. This last dataset also shows that non-perceptual aspects of decision-making can arbitrarily produce both positive and negative psychometric shifts, consistent with results produced by Morgan et al. (2011). Together, our experiments show that confidence estimates can be a valuable source of information for distinguishing perceptual aftereffects from criterion shifts. Our experiments also suggest that changes to sensory coding drive the classic motion aftereffect, but post-perceptual processes drive the implied motion aftereffect.

One point worth emphasising is the equal precision of our decision and confidence estimates. In all experiments, the precision of our confidence estimates (as measured by the full width at the half height of raised Gaussian function fits) was equivalent to the precision of decision estimates (as measured by distances between the 25 and 75% points of Cumulative Gaussian fits; see Table1). Our two core measures thus provided equally precise measures of behavioural and subjective metacognitive performance. These findings support the hypothesis that aftereffects can be driven independently by both sensory and cognitive processes.

Our experiments address a contemporary debate, about whether perception is directly penetrated by cognition (in particular see Firestone & Scholl, 2016; Pylyshyn, 1999). Aftereffects induced by implied or imagined motion have been interpreted as top-down effects of cognition on motion perception (Winawer et al., 2008, 2010; Pavan et al., 2011). This claim is bolstered by physiological

evidence that viewing photographs depicting implied motion activates some of the same direction-selective cortical circuits as viewing real motion, as measured by EEG (Lorteije et al., 2006; Lorteije et al., 2007). It is not surprising, however, that motion-related cognitions rely on some of the same neural circuits as perceiving motion. Indeed, it would be more surprising if there were no neural overlap between thinking about and perceiving motion. Our data simply suggest that activation of motion circuitry is not the same as perceiving motion. Whereas adaptation to real motion uniformly impacts decision and confidence reports, the implied motion aftereffect selectively impacts direction decisions under uncertainty. Since both perceptual decisions and confidence reports are informed by sensory evidence, this pattern of results implicates post-perceptual processes as driving the implied motion aftereffect, not sensory adaptation.

Our data cannot determine why implied motion adaptation produces a systematic negative aftereffect, which some have argued is indicative of a perceptual aftereffect (Fritsche et al., 2017; but note that this suggestion contrasts with our Experiment 3 findings). Our core argument, however, is that an aftereffect is comprised of either perceptual or post-perceptual changes, or both, and these factors can produce additive behavioural effects in the categorisation tasks typically used to measure aftereffects. Also, our data demonstrate that perceptual confidence is less impacted by post-perceptual processes relative to categorical decisions. Future experiments could therefore use confidence to estimate the composition of aftereffects, clarifying whether they have a perceptual causal locus (see Storrs, 2015).

In conclusion, measuring decision changes before and after adaptation often provides insufficient evidence to determine if an aftereffect is perceptual in origin. Confidence judgements can provide important additional diagnostic information. We have advanced a simple metacognitive approach to determining

whether any particular aftereffect makes the world look (or sound, or feel) different.

Authorship Declaration

R. M. Gallagher and D. H. Arnold developed the concept and experimental designs. R. M. Gallagher tested participants and performed statistical analyses. R. M. Gallagher drafted the manuscript under the supervision of T. Suddendorf, & D. H. Arnold, who provided critical revisions. All authors wrote and approved the final version of the manuscript.

References

Anstis, S. M. (2002). Was El Greco astigmatic? *Leonardo*, *35*(2), 208-208.

Banerjee, P., Chatterjee, P., & Sinha, J. (2012). Is it light or dark? Recalling moral behavior changes perception of brightness. *Psychological Science*, *23*(4), 407-409.

Barlow, H. B., & Hill, R. M. (1963). Evidence for a physiological explanation of the waterfall phenomenon and figural after-effects. *Nature*, *200*(4913), 1345-1347.

Brainard, D. H. (1997). The psychophysics toolbox. Spatial Vision, 10, 433–436.

Clifford, C. W., Webster, M. A., Stanley, G. B., Stocker, A. A., Kohn, A., Sharpee, T. O., & Schwartz, O. (2007). Visual adaptation: Neural, psychological and computational aspects. *Vision research*, *47*(25), 3125-3131.

Firestone, C., & Scholl, B. J. (2014). "Top-down" effects where none should be found the El Greco fallacy in perception research. *Psychological science*, *25*(1), 38-46.

Firestone, C., & Scholl, B. J. (2016). Cognition does not affect perception: Evaluating the evidence for" top-down" effects. Behavioral and brain sciences, 39.

Fritsche, M., Mostert, P., & de Lange, F. P. (2017). Opposite effects of recent history on perception and decision. Current Biology, 27(4), 590-595.

Lorteije, J.A.M., Kenemans, J.L., Jellema, T., van der Lubbe, R.H.J., de Heer, F., & van Wezel, R.J.A. (2006). Delayed response to animate implied motion in human motion processing areas. Journal of Cognitive Neuroscience, 18, 158–168.

Lorteije, J.A.M., Kenemans, J.L., Jellema, T., van der Lubbe, R.H.J., Lommers, M.W., & van Wezel, R.J. (2007). Adaptation to real motion reveals direction-

selective interactions between real and implied motion processing. Journal of Cognitive Neuroscience, 19, 1231–1240.

Mollon, J. (1974). After-effects and the brain. *New Scientist*, *61*(886), 479-482.

Morgan M., Dillenburger B., Raphael S., Solomon J. A. (2011). Observers can voluntarily shift their psychometric functions without losing sensitivity. Attention Perception & Psychophysics, 74, 185 – 193.

Pavan, A., Cuturi, L. F., Maniglia, M., Casco, C., & Campana, G. (2011). Implied motion from static photographs influences the perceived position of stationary objects. Vision research, 51(1), 187-194.

Pelli D. G. (1997). The Video Toolbox software for visual psychophysics: Transforming numbers into movies. Spatial Vision, 10, 437–442.

Pylyshyn, Z. (1999). Is vision continuous with cognition?: The case for cognitive impenetrability of visual perception. Behavioral and brain sciences, 22(3), 341-365.

Storrs, K. R. (2015). Are high-level aftereffects perceptual? Frontiers in psychology, 6.

Tanner Jr, W. P., & Swets, J. A. (1954). A decision-making theory of visual detection. *Psychological review*, *61*(6), 401.

Tversky, A., & Kahneman, D. (1971). Belief in the law of small numbers. Psychological bulletin, 76(2), 105.

Wark, B., Lundstrom, B. N., & Fairhall, A. (2007). Sensory adaptation. *Current opinion in neurobiology*, *17*(4), 423-429.

Winawer, J., Huk, A. C., & Boroditsky, L. (2008). A motion aftereffect from still photographs depicting motion. *Psychological Science*, *19*(3), 276-283.

Winawer, J., Huk, A. C., & Boroditsky, L. (2010). A motion aftereffect from visual imagery of motion. *Cognition*, *114*(2), 276-284.

Yarrow, K., Jahn, N., Durant, S., & Arnold, D. H. (2011). Shifts of criteria or neural timing? The assumptions underlying timing perception studies. Consciousness and cognition, 20(4), 1518-1531.