Illusory line motion in onset and offset bars

Sihang Han[1] · **Zhuoying Zhu**[1] · **Michael C. Corballis**[1] · **Jeff P. Hamm**[1]

Abstract Illusory line motion (ILM) refers to the perception of motion in a bar that is presented all at once next to an inducing stimulus. The experimental methods for producing and quantifying ILM are varied, and the resulting explanations are likewise at odds. The current study examined the explanations for ILM away from the inducing stimulus (bright or dark flash) using bars that either suddenly appear (onset bars) or suddenly disappear (offset bars). Real motion is used to cancel ILM, providing three measures to quantify ILM: the consistency of responding to only ILM in the absence of real motion; the distance between the points of subjective equality between ILM and the real motion; and the area between the curves relating perception of motion to the real and illusory conditions. ILM quantities for onset and offset bars are strongly correlated when the bar change occurs after the flash. However, onset bars presented during the flash do not show any evidence of ILM, whereas offset bars removed during the flash result in reverse ILM (rILM). Moreover, rILM and ILM are not correlated, suggesting they reflect two separate illusions. These results are consistent over the various measures of ILM.

Keywords Attention · Illusory line motion · Motion perception

Forms of illusory line motion (ILM) have been investigated since the early 1900s. There have been a number of paradigms employed, and the phenomenon has been quantified in different ways. As a result, a range of explanations has been offered. This variety, while providing a rich source of data, presents its own problems. Primarily, the lack of consistency in protocols and measurements makes it difficult to know if the results arising from one study are relevant to a study using different methods or measuring different quantities. One cannot be sure they are examining the same underlying phenomena. What follows is not so much an attempt to determine which explanations of illusory line motion are supported or valid and which can be dismissed or rejected, but rather a presentation that attempts to determine what findings warrant a common explanation and what findings can be argued to be reflective of a different illusory phenomenon. After coverage of the literature illustrating the variations of methods, we focused primarily on illusory line motion that occurs following a brief luminance change and made recommendations with respect to experimental methods, quantification of ILM, and data analyses.

Brief overview of illusory line motion

According to Kanizsa (1979), F. Kenkel in 1913 coined the term "gamma motion" to refer to the motion that is perceived when an object is suddenly presented. Kenkel described gamma motion in terms of the object expanding from its centre of mass (Faubert & von Grünau, 1995). Kanizsa demonstrated that if the object appeared next to an existing object, then the gamma motion was directed away from the preexisting object (Kanizsa, 1979). This is called polarized gamma motion (Kanizsa, 1979), although in similar displays von Grünau and Faubert (1994) referred to this as motion induction. Although the presented object may be of any shape (Kanizsa, 1979), for convenience and to coincide with the

✉ Jeff P. Hamm
j.hamm@auckland.ac.nz

[1] Centre for Brain Research, School of Psychology, The University of Auckland, City Campus, Private Bag 92019, Auckland, New Zealand 1142

current study and recent literature (Christie & Barresi, 2002; Christie & Klein, 2005; Crawford et al., 2010; Crawford, Kean, Klein, & Hamm, 2006; Downing & Treisman, 1997; Hamm et al., 2014; Hamm & Klein, 2002; Hikosaka, Miyauchi, & Shimojo, 1993a, 1993b, 1993c; von Grünau & Faubert, 1994), the object that undergoes illusory motion is referred to as the bar, because this object is generally rectangular with the motion directed along the long axis. The existing object that induces the illusory motion in the bar is referred to as the box, because the inducing object is typically a solid square. When referring to the phenomenon rather than the stimulus components of the display, the term "line" is used rather than bar, as in illusory line motion or ILM.

When the bar is suddenly presented next to a single preexisting box, the percept is of the bar emerging from the box. However, if an existing bar is removed, it appears to be sucked into the box (Kanizsa, 1979). A simple explanation for these single-box-and-bar findings could be that motion detectors are responding to a change in the location of the centre of mass. When only the box is presented, the centre of mass is located within the box, and when the bar appears the centre of mass shifts to within the bar. Similarly, when an existing bar is removed, the centre of mass shifts into the remaining box; the result is perception of motion towards the box (Zanker, 1997). Alternatively, motion has been suggested to occur away from a contrast decrease and towards a simultaneous contrast increase (Hock & Nichols, 2010). When the bar onsets, the motion is away from the box and the edge between the box and bar vanishes and provides the contrast decrease while the bar's surface and far edge create the contrast increase with respect to the background. When the bar is offset, the motion is towards the box; here, the removal of the far edge and the bar surface produce contrast decrease while the revealing of the edge of the box provides the contrast increase.

However, contrast counter-change cannot account for onset bars appearing to move away from an existing item even if the bar does not touch the item, and so there is no contrast decrease, because the bar is simply closer to one item than others in the display (von Grünau, Dube, & Kwas, 1996). Moreover, if a bar appears that is bright at one end and progressively dims along its length, motion occurs from the bright to the dim end (von Grünau, Racette, & Kwas, 1996); this occurs despite the fact that the entire bar presentation creates only a contrast increase. Furthermore, the motion is in the direction from the larger increase towards the smaller, so it is towards a tendency to decrease. This finding cannot be accounted for by a shift in the centre of mass either as the entire bar is presented simultaneously. Finally, motion can occur in an onset bar in an otherwise blank field following nonvisual cuing (Shimojo, Miyauchi, & Hikosaka, 1997) or through reflexive gaze orienting of attention (Bavelier, Schneider, & Monacelli, 2002) to the location where one end of the bar

will appear. Again, these displays contain neither a contrast decrease nor a shift in the centre of mass.

Moreover, von Grünau and Faubert (1994) have demonstrated that the motion does not require the boxes and bars to be defined by luminance but may be defined by a number of nonluminance-based attributes (i.e., motion, stereo-depth, etc). They suggest that the illusory motion is due to attention being directed to the box, speeding perception of the bar at the end closest to the box and creating the perception of movement away from the box. They call this attentional priming, although there appears to be a nonattentional process that produces illusory motion as well (von Grünau, Dube, et al., 1996). A further challenge to theories of ILM comes from the case when the bar is presented centred between two boxes forming a barbell type display; the motion percept is described as the bar crashing in the middle (Faubert & von Grünau, 1995). While this motion can be explained by contrast counter-change, it cannot be attributed to a change in the centre of mass, which remains midway between the boxes upon the presentation of the bar. Faubert and von Grünau (1995) refer to this as split priming.

When one of the two boxes flashes prior to the bar presentation then the motion is away from the flash rather than towards the middle. Visual attention has been used to explain this motion (Hikosaka et al., 1993a, 1993b, 1993c), because the flash exogenously attracts attention (Hamm & Klein, 2002; Posner & Cohen, 1984) resulting in prior entry of the end near the flash so that it is perceived sooner than the far end, creating the signal of a bar in motion.

ILM is thought to result from exogenously oriented attention but not endogenously oriented attention (Christie, 2014; Christie & Klein, 2005), which is consistent with the larger prior entry benefits associated with exogenous attention over that of endogenous attention (Shore, Spence, & Klein, 2001). These prior entry benefits of attention are thought to reduce as a function of distance (Laberge, 1983), resulting in a gradient of perceptual onsets that replicate the low-level signal of a bar in motion (Hikosaka et al., 1993a, 1993b, 1993c). Moreover, exogenously captured attention has been shown to be associated with the conscious perception of near threshold targets (Chica, Lasaponara, Lupianez, Doricchi, & Bartolomeo, 2010) further indicating that the prior-entry benefits are not simply a response bias but actually influence the conscious perception time of stimuli. Importantly, when the term attention is used within this study it should be taken to refer only to exogenous attention as endogenous attention appears not to produce ILM (Christie, 2014; Christie & Klein, 2005).

When a preexisting bar is removed from between two boxes following the flash, the illusory motion of the disappearing bar is also away from the flashed box (Crawford et al., 2010; Hamm et al., 2014). This is unlike polarized gamma motion where the motion is towards the single box when an existing bar is removed in the absence of a flash (Kanizsa, 1979); there

may be separate underlying causes behind polarized gamma motion and illusory line motion away from a flash. Indeed, the possibility that there may be multiple different forms of illusory line motion is further developed in von Grünau, Dube, and Kwas (1996), where it is suggested that ILM may arise from preattentive processes as well as from an attentional gradient.

Consistent with the attentional aspects of the explanation for ILM are the findings that the neural networks associated with visual attention are activated by these two boxes, flash, and bar type displays (Hamm et al., 2014; Tanabe & Yanagida, 2002). Moreover, ILM has been shown to be weaker in people diagnosed with schizophrenia (Crawford et al., 2010), a disorder long known to be associated with deficits in visual attention (Kraepelin, 1919). Additionally, illusory line motion is reported even in the absence of a box, such as following reflexive gaze orienting (Bavelier et al., 2002) or when attention is direction to a location by auditory or tactile cues (Shimojo et al., 1997). Finally, illusory line motion in these displays can influence saccadic eye movements to the point of reversing the pro-saccade advantage by speeding eye-movements away from the flashed location (Crawford et al., 2006).

A major requirement of the attentional gradient explanation is that attention must be focused at a location prior to the onset/offset of the bar stimulus. Polarised gamma motion, however, simply requires that the bar is presented or removed in the vicinity of an existing stimulus even in situations where attention is not directed to that location (von Grünau, Dube, et al., 1996). Finally, gamma motion occurs even when an object singularly appears or disappears in an otherwise blank visual field (Kenkel as cited in Faubert & von Grünau, 1995) and seems to reflect the speeded detection or prolonged retention of the object's surface at the centre of mass. In other words, even if exogenous attention does result in ILM it cannot be the sole cause as additional explanations would be required for polarized gamma motion and gamma motion.

However, there is no reason why the illusion of motion that occurs in these different displays must be attributed to a single cause. For example, gamma motion could arise due to a gradient of increased activity that spreads from the centre of mass of an object. Polarised gamma motion may reflect the subthreshold spreading of activation produced by visual stimuli simply because they are present (Jancke, Chavane, Naarman, & Girinvald, 2004) and/or motion may be signalled by simultaneous opposite changes in contrast (Hock & Nichols, 2010) or by the change in location of the centre of mass (Zanker, 1997). Finally, a flash may result in temporary prior entry benefits that arise at a given location when exogenous attention has been captured by a salient event at this location (Shore et al., 2001). There is nothing about these explanations that prohibits all of them occurring simultaneously; they are not mutually exclusive.

Furthering the notion that there may be multiple routes to ILM is the fact that some demonstrations of ILM appear to be difficult to explain other than by a gradient of attention (Bavelier et al., 2002; Shimojo et al., 1997) and yet there also are demonstrations that cannot be explained by a preexisting gradient of attention. As an example of the latter, it has been shown that if the two boxes are in different colours, then upon presentation of a coloured bar that matches the colour of one of the boxes, illusory motion is away from the matching box (Faubert & von Grünau, 1995). Similarly, Hock and Nichols (2010) presented boxes of different luminance rather than hue that were joined by a bar of intermediate luminance, which then changed luminance to match either one or the other box. The result was ILM away from the box of similar luminance. If the starting boxes are different sizes, one tall one short, and the bar that appears matches the height of one of the boxes then the motion will be away from the box whose height matches the height of the bar (Corballis, Funnell, & Gazzaniga, 2002). The motion in these displays is referred to as arising due to attribute priming. It is clear that ILM as a result of attribute priming cannot be due to a bias in the focus of attention prior to the presentation of the bar nor is the size based version easy to explain in terms of the spreading of subthreshold activation as one would expect this always to be greater in the vicinity of the taller box.

Also difficult to explain by a gradient of attention is the reverse ILM that occurs when a bar is removed during the flash. If the bar is removed after the flash is over the motion is away from the flash (Crawford et al., 2010; Hamm et al., 2014); however, if the bar is removed during the flash then the motion is towards the flash (Hamm et al., 2014); this is reverse ILM (rILM). While some have suggested that attention may result in prolonging existing visual signals (Schmidt & Klein, 1997), which would bring rILM under an attentional explanation, it does not explain why offset bars produce rILM during the flash but ILM after the flash. Rather, rILM could reflect a different cause, such as spreading of subthreshold activations outward from the brighter box prolonging the stimulus near the box after offset (Jancke et al., 2004) such that the far end disappears before the near end.

ILM based on attribute priming could be explained as reflecting a reinterpretation of the display such that the onset of the bar is linked or associated with the same coloured box. This type of explanation would coincide with Downing and Treisman's (1997) impletion theory of ILM which holds that the motion is a result of the reinterpretation an implausible visual event into a more plausible, real-world, event. Effectively, bars do not suddenly appear out of nowhere and so the motion percept is added to increase the plausibility of the display. If this were the case, then it is reasonable to assume that a luminance increase might be more strongly associated with a bar onset than offset, whereas a luminance decrease would be more strongly associated with a bar offset.

The impletion explanation risks becoming circular—whatever situation results in motion becomes deemed to be the more plausible real-world interpretation as evidenced by the perception of the illusory motion.

It appears, therefore, that in the literature ILM may refer to as many as four separate illusions of motion: 1) gamma motion (Kanizsa, 1979) that could reflect a gradient of activity radiating from the centre of mass of an object; 2) polarized gamma motion, which could reflect shifts of the centre of mass (Zanker, 1997), contrast counter-change (Hock & Nichols, 2010), and/or increased activity that spreads out from a visual stimulus (Jancke et al., 2004) resulting in faster detection of nearby onset stimuli and prolonged perception of offset stimuli; 3) motion away from a box when the bar matches a physical attribute (colour, luminance, size) of the box (Corballis et al., 2002; Faubert & von Grünau, 1995; Hock & Nichols, 2010), referred to as attribute priming, which may reflect contrast counter-change; and 4) motion away from a flashed box towards a nonflashed box, which may reflect an attentional gradient of prior entry benefits centred on the flash (Hamm et al., 2014; Hikosaka et al., 1993a, 1993b, 1993c). As a collection, these may describe how the visual system interprets incoming visual information to maximize the detection of actually moving stimuli. In other words, these may be the processes by which impletion (Downing & Treisman, 1997) is achieved without invoking a circular explanation.

In short, there may be a number of different ways to achieve an illusion of motion. Therefore, when testing explanations for why these illusions occur, it is important to ensure that changes in the display conditions do not result in illusions of motion for different reasons. For example, while the attentional gradient model can easily explain, and in fact predicts, ILM following a brief peripheral flash it cannot explain the luminance, colour, or size attribute priming based ILM (Corballis et al., 2002; Faubert & von Grünau, 1995; Hock & Nichols, 2010). However, explanations for attribute priming are based on visual interactions between the matching box and bar and cannot account for cross-model cuing (Hikosaka, Miyauchi, Takeuchi, & Shimojo, 1996; Shimojo et al., 1997) or reflexive gaze orienting (Bavelier et al., 2002), resulting in ILM, because there is no box with which the bar can correspond.

If, therefore, there are multiple forms of ILM being activated by the different display protocols, it becomes difficult to test theoretical explanations of ILM if one cannot be sure if two protocols are testing or eliciting the same illusion. The purpose of the current study was to present a correlational methodology based on an individual differences approach as a basis for determining if two experimental protocols result in the same illusion and therefore require theory to account for both sets of findings. If, however, the two protocols appear to result in illusions that are unrelated other than to be illusions of motion, then there is no inherent need for a theoretical

explanation that accounts for one to be concerned with the other. In fact, in this latter case, a single theory that accounts for both illusions would be beholden to explain why the data suggests the illusions are unrelated.

It is beyond the scope of this study to test all display configurations; rather it is the goal to provide a general approach towards addressing the issue. Specifically, the current study was designed to determine whether illusions arising during two boxes, flash, and bar type experiments are the same illusion and so require a common explanation. In Experiment 1, the specific question was whether or not the same illusion is being invoked in the four combinations created from brightening or dimming one of the boxes prior to the bar either appearing or disappearing.

It has been shown that if the bar is removed during the flash rather than after it that the illusory motion is towards the flash and not away from the flash (Hamm et al., 2014). Experiments 2 and 3 examined whether this reverse ILM (rILM) arose with onset bars (Experiment 2) and with offset bars (Experiment 3) and examined whether this rILM was the same illusion and therefore necessitated a common explanation with the illusion that occurred when the bar was changed after the flash and resulted in ILM away from the flash.

While individual studies have shown that the illusory motion is away from a bright flash for both onset (Hikosaka et al., 1993c) and offset (Crawford et al., 2010; Hamm et al., 2014) bars, so far it has not been determined whether these illusions are correlated at the individual level. It is possible that onset and offset bars result in illusory motion for entirely different reasons. If onset and offset illusions arose for unrelated reasons, there would be no reason to expect that a person who showed a large illusion for onset bars also showed a large illusion for offset bars. If both illusions arose for a common reason, then they should be correlated at the individual level, meaning that an individual who showed a large illusion with onset bars should likewise show a large illusion with offset bars.

In summary, the major purpose of Experiment 1 was to examine ILM in the context of onset and offset bars that are preceded by a brightening or diming of one of the two boxes. Moreover, the goal was to determine if the various combinations result in illusions of motion that are likely to arise due to a common underlying mechanism. This question is addressed by looking at the correlation between the ILM obtained in the various conditions at the individual participant level.

Approach to data analysis and interpretation

Throughout all experiments presented, decisions were made based on the accuracy of the predictions derived from the null hypothesis. A relationship between two conditions or an effect of a manipulation was only considered necessary if the

obtained data were deemed improbable to occur due to natural variability alone, which is the prediction derived from the null hypothesis. As is typical in psychological research, the criterion for rejecting the null hypothesis is set at $p < 0.05$. Because multiple attempts to reject the same null hypothesis results in the calculated p value underestimating the actual probability of obtaining the observed data if the null hypothesis is accurate, Bonferroni correction was employed when appropriate. While not wishing to be ruled by the p value, unless there was a compelling reason to contradict the objective criterions set out here, the analysis will be discussed along the lines indicated by the analysis.

Whereas a significant correlation would suggest that two conditions are related, there is always concern about concluding that a nonsignificant correlation reflects the absence of a relationship. However, because the lack of a relationship is of theoretical importance, the relative support for accepting the null hypothesis will be assessed by calculating the Bayesian probability of the null hypothesis given the data, or $p_{H0|D}$ (Masson, 2011). These values will be described using labels suggested by Raftery (see Table 6, Raftery, 1995), such that $p_{H0|D}$ of 0.5-0.75 is weak evidence *in favour* of the null hypothesis, 0.75^{+}-0.95 is considered positive evidence *in favour* of the null hypothesis, 0.95^{+}-0.99 is considered strong evidence *in favour* of the null hypothesis, and >0.99 is very strong evidence *in favour* of the null hypothesis. To avoid confusion by switching between $p_{H0|D}$ and $p_{H1|D}$, when evidence is in favour of the alternative hypothesis, the Bayes probability will still be discussed in terms of the probability of the null hypothesis, such that values of 0.50^{-}-0.25 is considered weak evidence *against* the null hypothesis, and 0.25^{-}-0.05 as positive evidence *against* the null hypothesis, 0.05^{-}-0.01 is strong evidence *against* the null hypothesis, and <0.01 is very strong evidence *against* the null hypothesis.

Quantification of ILM

Determining whether the illusion that occurs from two different display conditions is correlated is much simpler if we quantify the illusion by a single value. The field of ILM research is not only an embarrassment of riches when it comes to display protocols that produce illusory motion, but there is an overabundance of different approaches for how to measure it. For example, in some studies the illusion was scored simply as the percentage of responses that indicate motion away from the box (Corballis et al., 2002; Hikosaka et al., 1993c; von Grünau & Faubert, 1994). Sometimes a subjective rating scale was used, where participants were asked to rate the strength of their motion perception (Christie, 2014; Christie & Klein, 2005). Luminance gratings (von Grünau, Racette, et al., 1996) have been used to produce presumed perceptual asymmetries resulting in motion in the opposite direction until

a point of subjective equality is found, quantifying the illusion in terms of the luminance grating. Similarly, the bar has been presented in sections, towards or away from the flashed box, to quantify the illusion in terms of the amount of real motion required to reach the point of subjective equality (Bavelier et al., 2002; Crawford et al., 2010; Steinman, Steinman, & Lehmkuhle, 1995). This last method was used in the current study because it produces a reliable pattern of results that allowed for the collection of three perceptual metrics for ILM.

The use of real motion to cancel illusory motion was achieved by presenting portions of the bar over successive screen refreshes on the computer screen. The no real motion condition involves presenting the bar in its entirety on a single screen refresh. The fastest motion is achieved when the bar is presented in halves over two screens. Medium speed presents the bar in thirds over three screens. The slowest motion involves presenting the bar in quarters over four screens. These motions can be coded as having 0, 1, 2, or 3 screen refreshes between the left and right edge being presented. Because the real motion can be leftward or rightward, direction was coded by using negative values to indicate leftward motion; so, the conditions are coded from −3 (slowest leftward) to +3 (slowest rightward).

Motion perception was indicated by a two-alternative forced choice between leftward or rightward motion. When the data are scored in terms of proportion of rightward responses the design provides data for three psychometric functions, one resulting from each of the three possible flash conditions, namely left, right, or no flash. The percentage rightward scores can be converted to a percept score, by multiplying by two and then subtracting one, which is the same as coding left responses as −1 and right responses as +1 and averaging over the responses made. The mean percept scores range from −1 to +1 rather than between 0 and 1. The benefit of this range is that negative scores indicate the condition is perceived as more often to the left than right, positive scores indicate more often to the right than left, and a score of 0 indicates no consistent motion direction was perceived. This corresponds to negative values indicating leftward real motion and positive values indicating real rightward motion and so negative and positive values similarly indicate left and right on both the x and y axes. A stylised figure illustrating the type of data obtained from the paradigm described is shown in Fig. 1.

ILM can be quantified as the difference between the percept scores following a left and right flash when there is no real motion in the display, as represented by the distance between A and B in Fig. 1. This measure should always be calculated as the value from the right flash condition (B) subtracted from the left flash condition (A), because if the motion were towards the flash the calculated value is negative. This maintains a change in sign as indicating a change in direction, but it also ensures that conditions that do not produce an illusion can have values that can vary either side of
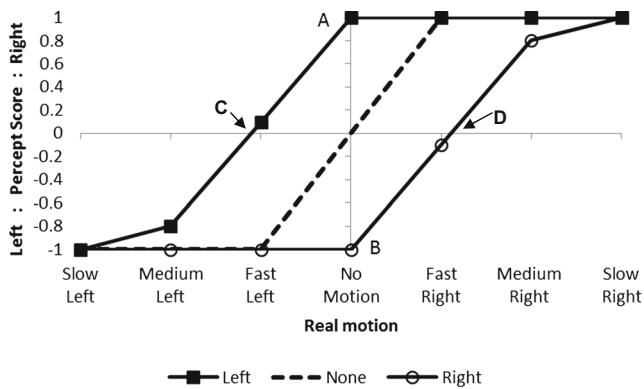
Fig. 1 Stylised figure illustrating percept scores as a function of real motion when a flash is presented before the bar at the box on the left (solid line with filled squares), right (solid line with open circles), or neither (dashed line with asterisks). ILM$_{area}$ represents the area between the curves for the left and right flash conditions. ILM$_{con}$ is the difference between the percept scores indicated by A and B. ILM$_{PSE}$ is derived after fitting a logistic function to the data to obtain the points of subjectivity and determining the distance between them, as indicated by C and D

zero. As a result, the range of values is between −2 and +2. This measure effectively reflects the consistency of the illusion in the absence of real motion and will be referred to as ILM$_{con}$. This measure corresponds to what is collected in many studies as illusory line motion often is studied in experimental conditions that do not include any trials on which real motion is presented.

ILM also can be quantified as the distance between the points of subjective equality (PSE), referred to as ILM$_{PSE}$, and is represented as the distance between the points where the data curves cross the x-axis as indicated by C and D in Fig. 1. This measure is calculated as the PSE for the left flash condition subtracted from the PSE for the right flash condition, which again results in a negative value if the illusion is towards the flash. Because it is possible for the PSE to be beyond the points measured on the x-axis, ILM$_{PSE}$ could theoretically take on any value between −∞ and +∞.

ILM$_{PSE}$ as a measure combines the speed of leftward and rightward motion and would reflect the combined speed at which bars under each of the two illusions appear to move apart from each other. It is possible that an experimental manipulation could influence the speed of the perceived illusion and yet not influence the consistency with which participants report the direction of that motion. Therefore, making comparisons between findings in the literature can be difficult as the degree to which ILM$_{con}$ and ILM$_{PSE}$ are related is not established. There is good reason to be concerned, because the measures are theoretically orthogonal dimensions (as illustrated in Fig. 1).

Finally, the third metric quantifies ILM as the area between the curves for the left and right flash conditions, which is obtained by calculating a series of trapezoids based on the observed data for the left and right data functions and then obtaining the difference by subtracting the latter from the former. This will be referred to as ILM$_{area}$. Again, because the area under the right flash curve is subtracted from the area under the left flash curve to produce this metric, ILM towards the flash produces a negative value and no illusion is represented by a value of zero. Given the perceptual space being tested in the current study the values would be bounded between −12 and +12, but these boundary values are specific to the study in question.

To summarize, the ILM$_{PSE}$ measure is suggested to index the speed of the illusion. The ILM$_{con}$ measure reflects the consistency of the response when there is no real motion and would index the saliency of the illusion. The area measure, which combines speed and saliency, is suggested as an overall measure of the size of the illusion. Each of these measures has its own pros and cons. Whereas ILM$_{con}$ could be criticized for being based on a limited subset of the data, it is derived solely from the data from the no real motion condition and so any perceived motion can only be illusory motion.

ILM$_{PSE}$ has the benefit of being derived from a function that is fit to all of the collected data. However, because it involves fitting a function to the data before then extracting the measure it becomes that much further removed from the actual observations. This becomes an increasing problem the worse the fit of the psychometric function to the observed data. Moreover, there are various choices of functions that one could choose to model the data, log linear functions or Weibull functions for example, each of which would result in slightly different values. Moreover, there will be some degree of error between the data and the curve fit; this introduces a source of error in the PSE measure that is not part of the observations but is based on the theoretical curve.

Similar to ILM$_{PSE}$, ILM$_{area}$ is calculated using all of the observations, and similar to ILM$_{con}$, it is calculated directly from the observed data and does not require the intermediate step of fitting a theoretical curve. However, while ILM$_{PSE}$ and ILM$_{con}$ are theoretically orthogonal measures and could be considered capturing specific aspects of the motion perception, such as speed and saliency, ILM$_{area}$ are influenced by the speed or saliency of the illusion, or both. Being a combined measure means that caution is required to avoid making unduly specific claims about characteristics of the illusion percept, but it is a benefit if the question of investigation is not specific to issues of speed or saliency. This combined measure could be said to reflect the illusion's overall size or magnitude.

While it is tempting to look for left and right biases by calculating changes relative to the no flash condition, such as ILM$_{area(left)}$ − ILM$_{area(none)}$ compared with ILM$_{area(none)}$ − ILM$_{area(right)}$, this suffers from similar problems associated with attempting to calculate separate costs and benefits rather than simply the combined costs plus benefits (Jonides & Mack, 1984). Therefore, while the no flash condition will be included in the experiments, and the resulting data presented

in the figures, the focus of the current study and analysis is on the values derived from the left and right flash conditions.

In addition to these perception based measures, the mean time to make the decision also is recorded. Decision times will reflect a wide range of processes other than just the perception of motion, including response selection, movement time, and response biases (Jensen, 2006). Because we are interested in the subjective perception of motion in a given condition, and not the accuracy of detecting the real motion that the condition contains, the decision time is the mean of all responses. The decision times tend to be slow when real motion and ILM conflict in direction and fast when real motion and ILM correspond in direction. No flash trials, so trials with no ILM, tend to have intermediate decision times (Crawford et al., 2010). Crawford et al. (2010) noted that the decision times tended to be slowest close to the point of subjective equality and noted how this indicated that the ILM motion signal and the real motion signal were combining to cancel each other rather than the PSE, indicating a point where there were two equally salient motion signals competing in a winner-take-all type situation. This relationship between the decision times and the PSE will be explored more fully.

In summary, Experiment 1 presented participants with bars that move to the left or right between two boxes; some trials presented the bar with no motion. Before the presentation of the bar one of the boxes may flash, which is expected to produce illusory motion away from the flashed side in both onset and offset bars. The illusions of motion in this display set up are believed to arise for similar underlying reasons. Three different measures of ILM were obtained to determine whether these measures are correlated with each other, which would aid in comparison between studies that quantify the illusion using different metrics. Finally, the notion that all of these display configurations, bright and dim cues combined with onset and offset bars, reflect a common underlying mechanism that results in the perception of illusory motion was tested by determining whether the measures are correlated between conditions at the individual participant level.

Methods

Participants

Twenty-four[1] naïve participants comprised of 13 females and 11 males between the ages of 18 and 40 years volunteered for the experiment. All were right-handed and had normal or adjusted-to-normal vision. Their mean age was 25.6 years. One female participant was dropped from the study due to near chance level discrimination during the offset bar conditions, resulting in usable data from 23 participants with a mean age of 25.8 years.

Apparatus

A 486 IBM-compatible personal computer, running custom software written in Borland Pascal 7.0 and executed under DOS 6.2 was used to control the presentation of stimuli and the collection of behavioural measures. All stimuli were displayed on a 14-inch colour monitor at a resolution of 640 × 480 pixels. Responses were made on the computer keypad on adjacent keys so left and right responses could be made with the index and middle finger of the right hand. The millisecond timing routines implemented are those described in Hamm (2001) and display synchronization with the monitor's refresh cycle was achieved as described by Heathcote (1988).

Stimuli

All stimuli were monochromatic on a grayscale of 64 levels with 0 being black and 63 white. The fixation point was a small ($0.5^0 \times 0.5^0$) black cross in the middle of the screen. Two light gray (grayscale = 40) square markers (2^0 per side) were displayed 1.2° above the cross, with their centres 5° to either side of the cross. The bar also was light gray (grayscale = 40), 8° in length and 1.5° in height, and displayed with its ends connecting to each of the two markers. All stimuli were presented against a neutral gray (grayscale = 30) background. The bright cue was defined as the sudden maximal increase in luminance (grayscale = 63 or white) of either square markers. The dim cue was the sudden maximal decrease in the marker luminance (grayscale = 0 or black).

Procedure

The experiment was conducted in a dimly lit room with each participant positioning their head on a chinrest 57 cm in front of the monitor. Participants were instructed to fixate the cross and after stimulus presentation to indicate the direction of the motion of the bar. Participants were told to choose the direction the line first appeared to move should it appear to start in one direction and then switch, as pilot testing of similar displays suggested was sometimes experienced. When no motion was detected, participants were asked to try and distribute guess responses equally on the two response keys and not to choose a default guess response. These instructions were given both verbally and on the computer screen before the commencement of the experiment.

---

[1] This number was chosen because the critical value that indicates the predictions from assuming no difference is r = 0.4044, which explains roughly 16 % of the variance, was deemed to be of practical value for general research purposes. This was especially so given that it was expected that if two conditions are resulting in the same illusion then, given the high similarity of the responses and decisions, the expected relationship is predicted to be quite high.

Design

Trials were divided into two blocks. In the onset bar block, trials began with the fixation cross and the two square markers for a fixation duration of 500 ms. The cue then appeared for 50 ms. Upon cue's offset (0 ms interstimulus interval (ISI)), the target line appeared providing a cue-line stimulus onset asynchrony (SOA) of 50 ms. Similar settings were used in the offset bar block, except that trials began with the bar connecting the two markers, and then disappearing after cue presentation. Presentations of the two blocks were counterbalanced with half the participants (n = 12) starting with the onset block and the other half (n = 12) with the offset block.

Within each block, the bar appeared or disappeared at one of seven real motion conditions: slow left, medium left, fast left, no motion, fast right, medium right, and slow right, which were coded as −3 through to +3, respectively. The speed of the real motion was manipulated by dividing the bar into segments of equal length, 4 segments for slow motion, 3 for medium motion, and 2 for fast motion. The segments appeared or disappeared on consecutive screen refreshes (16.67 msec per screen), either from left to right or right to left, to create real motion. The no-motion display was simply the entire bar appearing or disappearing on a single screen refresh. When coupled with the two different cue types (bright and dim) and three alternative cue locations (left, right, and none), there were in total 7 (motion; slow left, medium left, fast left, none, fast right, medium right, and slow right) × 2 (cue type; bright and dim) × 3 (cue location; left, none, and right) = 42 conditions. Within each block each condition was repeated 10 times, for a total of 420 trials in a block. Following a response there was a 1-sec intertrial interval. Every participant performed two blocks of trials during a single testing session, one with onset bars and one with offset bars, for a total of 840 trials. Figure 2 shows a depiction of a trial in which a bright left flash is followed by a slow leftward onset bar.

Decision times were measured from the initial point of the onset (or offset) of the bar, that is, when the first segment was presented (or removed) if real motion was involved, or the instant the entire bar was presented (or removed) if no real motion was involved. A left response was coded as −1 and a right response was coded as +1. The mean percept score refers to the average of these response codes and return values similar to a guess corrected accuracy measure.

Results

After discarding anticipation responses based on trials with a decision time less than 200 ms (0.93 % of the trials) and trials in which no response was made (0.16 %), the mean percept score for each cue brightness by line condition by cue location

were fitted to the following function; percept = $2[e^{ax + b}/(e^{ax+b} +1)] - 1$ by iterating the values for a and b to minimize the sum of the squared differences between the predicted scores and the observed data. This is fitting the same function as per a logistic regression $[e^{ax + b}/(e^{ax+b} +1)]$, except that the fit is based on least squares rather than maximum likelihood ratio. Points of subjective equality (PSE) were calculated by PSE = $-b/a$, which ranged between −4.61 and 7.26. The goodness of fit was expressed as a percentage reduction in the sum of squared error when calculated about the model as relative to when it was calculated about the mean. The average goodness of fit for a participant ranged between 71.3 % and 99.6 %.

The mean percept scores for the seven levels of real motion following a left, right, or no flash, for the various conditions can be seen in the upper half of Fig. 3 (A-E). Note, because no dim flash and no bright flash are identical conditions, data for these trials have been combined for both the percept scores and decision times. However, the no-flash condition is not required for the calculation of the quantities of interest, and so the duplication of the data shown in the figures is for display purposes only and does not affect any of the analyses. ILM can be quantified as the area between the curves for the left and right flash (ILM$_{area}$), as the horizontal distance between the points of subjective equality for the left and right flash conditions (ILM$_{PSE}$), or as the vertical distance between the percept scores for the left and right flash condition when there is no real motion (ILM$_{con}$).

ILM$_{area}$

The area between the curves was analysed in a two-way repeated measures ANOVA, with bar condition (onset and offset) and cue brightness (bright vs. dim) as factors. This resulted in a main effect of bar condition ($F_{(1,22)} = 4.69$, MSE = 1.207, $p = 0.042$, $\eta_p^2 = 0.176$; M = 6.39 vs. 5.89, onset vs. offset line condition, respectively), with no significant effect of cue brightness ($F_{(1,22)} = 1.71$, MSE = 0.402, $p = 0.205$, $\eta_p^2 = 0.072$; M = 6.23 vs. 6.05, for bright and dim cues, respectively), and no significant interaction ($F_{(1,22)} = 0.11$, MSE = 0.783, $p = 0.741$, $\eta_p^2 = 0.005$).

ILM$_{pse}$

The distance between the PSE was analysed in a similar two-way repeated measures ANOVA as ILM$_{area}$. This resulted in a main effect of bar condition ($F_{(1,22)} = 5.46$, MSE = 0.785, $p = 0.029$, $\eta_p^2 = 0.199$; M = 3.81 vs. 3.38, onset vs. offset bar condition, respectively), with no significant effect of cue brightness ($F_{(1,22)} = 1.51$, MSE = 0.967, $p = 0.231$, $\eta_p^2 = 0.065$; M = 3.72 vs. 3.46, for bright and dim cues, respectively), and no significant interaction ($F_{(1,22)} = 0.627$, MSE = 1.02, $p = 0.437$, $\eta_p^2 = 0.028$).

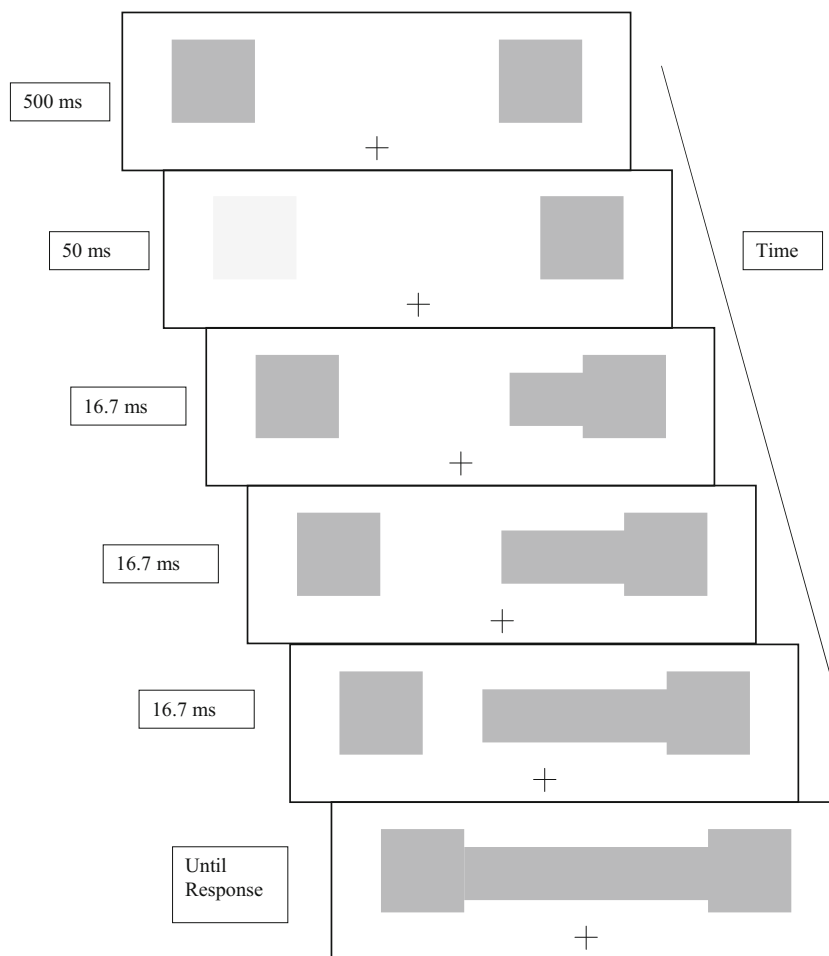**Fig. 2** Graphic depiction of a bright left flash trial followed by a slow left onset bar
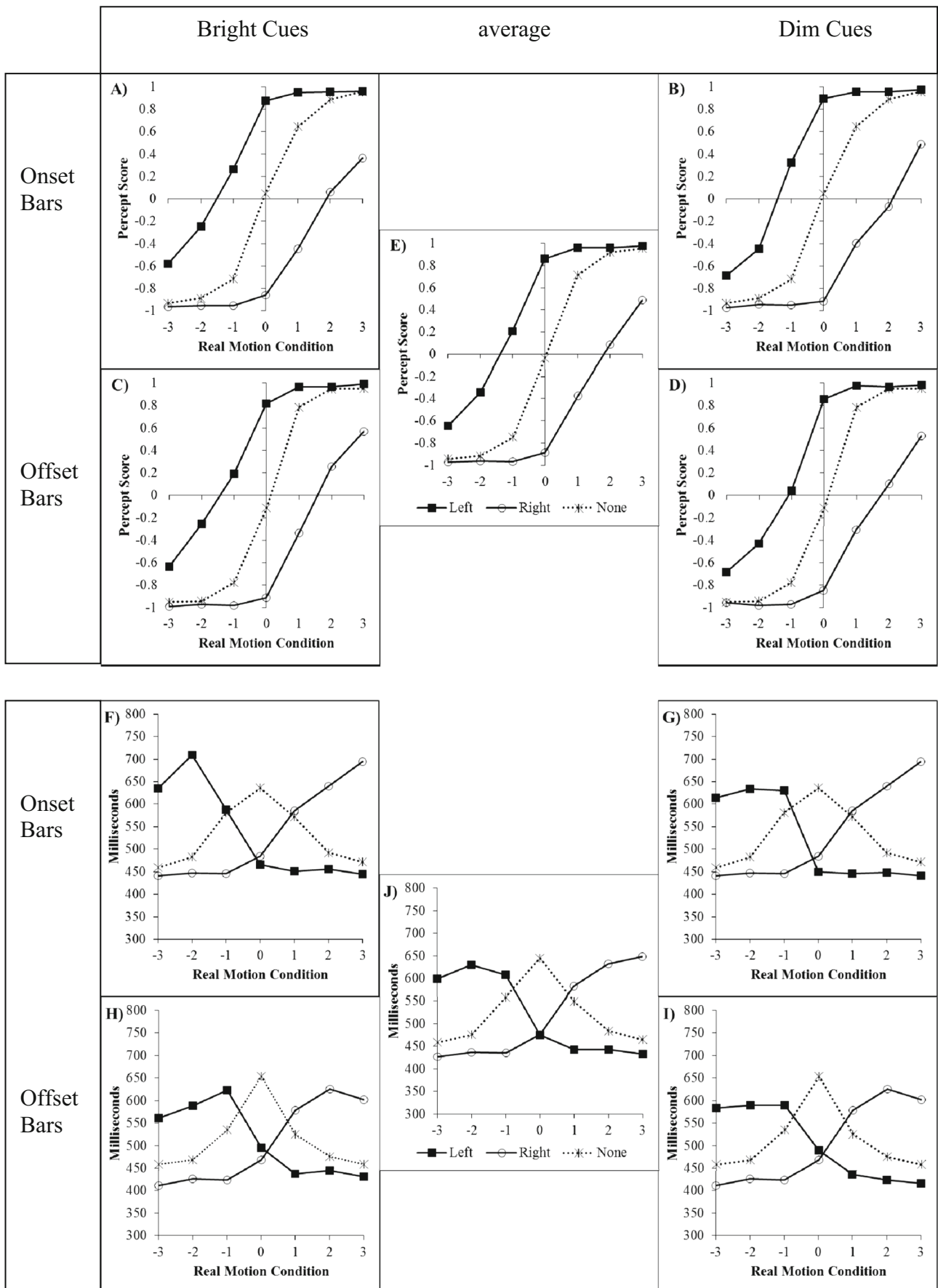
ILM$_{con}$

The distance between the percept scores when there was no real motion was analysed in a similar two-way repeated measures ANOVA as ILM$_{area}$ and ILM$_{PSE}$. This resulted in no main effect of bar condition ($F_{(1,22)} = 0.78$, MSE = 0.084, $p = 0.386$, $\eta_p^2 = 0.034$; M = 1.77 vs 1.72, onset vs. offset bar condition, respectively), with no significant effect of cue brightness ($F_{(1,22)} = 0.33$, MSE = 0.042, $p = 0.571$, $\eta_p^2 = 0.015$; M = 1.73 vs 1.76, for bright and dim cues, respectively), and no significant interaction ($F_{(1,22)} = 1.93$, MSE = 0.031, $p = 0.178$, $\eta_p^2 = 0.081$).

For all three measures of ILM, the data were averaged over cue brightness, and the measures for onset bars were correlated with the corresponding measure for offset bars to determine the test-retest reliability. Scatterplots showing these correlations are in Fig. 4. As each of these tests is effectively asking the same question via a different measure (is ILM for onset bars correlated with ILM for offset bars), a Bonferroni correction factor of 3 needs to be applied, making the criterion for significance $p < 0.0167$. Data pairs with excessive levels of influence on the correlation—as determined by having a Cook's distance > 0.1818 (4/(n-k-1), where n = 23 and k = 1)—were removed and are indicated by open symbols in the panels of Fig. 4. Due to the theoretical importance of a null relationship, the correlations were used to calculate the Bayesian probability of the null hypothesis given the data (Masson, 2011). Both ILM$_{area}$ (Fig. 4A) and ILM$_{PSE}$ (Fig. 4B) showed strong and significant relationships ($r_{(20)} = 0.81$, $p < 0.001$, $p_{(H0|D)} < 0.001$, and $r_{(18)} = 0.87$, $p < 0.001$, $p_{(H0|D)} < 0.001$) constituting very strong evidence against the null hypothesis. However, the relationship between ILM$_{con}$ (Fig. 4C) for onset and offset bars failed to reach significance, ($r_{(19)} = 0.42$, $p = 0.058$, $p_{(H0|D)} = 0.374$), and produced only weak evidence against the null hypothesis. The r values for ILM$_{area}$, ILM$_{PSE}$, and ILM$_{con}$ when the extreme values are included are 0.84, 0.83, and 0.39, with $p_{(H0|D)} < 0.001$, <0.001, and 0.418, respectively.

Each measure was then averaged across onset and offset bars, and the measures were correlated with each other as shown in Fig. 5, again removing data pairs with excessive Cook's distance values. There was very strong evidence against the null hypothesis produced by the significant correlation between the ILM$_{area}$ and ILM$_{PSE}$ (Fig. 5A; $r_{(20)} = 0.95$,

$p < 0.0001$, $p_{(H0|D)} < 0.001$), but neither ILM$_{area}$ (Fig. 5B; $r_{(20)}$ = 0.20, $p = 0.372$, $p_{(H0|D)} = 0.750$) nor ILM$_{PSE}$ (Fig. 5C; $r_{(19)}$ = 0.19, $p = 0.409$, $p_{(H0|D)} = 0.757$) measure correlated with ILM$_{con}$, resulting in weak and positive evidence in favour of the null hypothesis, respectively. When all data pairs are included the r values become, 0.88, 0.35, and 0.06, with $p_{(H0|D)}$ < 0.001, 0.516, and 0.822, respectively.

Decision times

The mean decision times are shown in the lower half of Fig. 3 (F-J). The group average mean decision times for all conditions were then plotted as a function of their distance from their point of subjective equality and are shown in Fig. 6. Note, although the no bright and no dim flash conditions are the same and the data are simply duplicated in Fig. 3, the values are only entered once into the distance decay function. This revealed a clear relationship, with response times tending to increase as the condition approached the point of subjective equality. The relationship was described as an exponential decay towards a lower asymptotic decision time value and the data was fitted to an equation of the form $y = dt_a + dt_0 \, e^{-rx}$, where dt_a is the decision time towards which the function asymptotically decays, dt_0 is the increase over dt_a at the PSE, the r exponent

indexes the rate of exponential decay, and x reflects the distance at which the decision time is to be predicted. The best fitting equation for these data was $dt = 383.7212 + 289.2056e^{-0.4574x}$, $r_{(77)} = 0.87$, $p < 0.001$, $p_{H0|D} < 0.001$, which is very strong evidence against the null hypothesis.

An additional exploratory measure, referred to as the congruency effect, also was calculated from the decision times, the analysis of which may be found in Appendix A.

Discussion

Three measures of ILM were calculated from the percept scores obtained during a cancelation procedure, namely the area between the left and right flash curves (ILM$_{area}$), the distance between the points of subjective equality (ILM$_{PSE}$) of the left and right flash curves, and the distance between the percept scores after a left and right flash in the absence of any real motion (ILM$_{con}$). The group mean decision times to make the motion judgement was analysed as a function of the distance from the group average points of subjective equality. These measures were obtained for both onset and offset bars. In addition, ILM could be induced by either a brightening or a dimming of the left or right peripheral box.

The ILM$_{area}$ and ILM$_{PSE}$ measures were both strongly correlated for both onset and offset bars for both the bright and dark cues resulting in very strong evidence against the null hypothesis of no relationship between ILM for onset and offset bars. This good test-retest reliability suggests that these displays all generate the same illusion, so any explanation for ILM needs to be able to account for all of these results. Moreover,
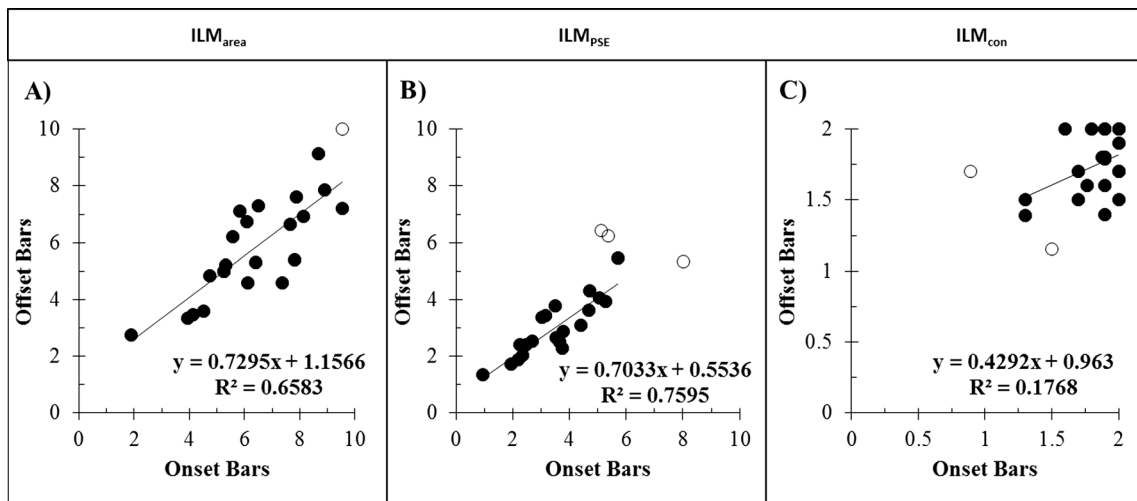

Fig. 4 Correlation between (A) ILM$_{area}$, (B) ILM$_{PSE}$, and (C) ILM$_{con}$ from the bar onset and bar offset conditions after collapsing over bright and dim cue conditions. Open circles indicate data pairs omitted due to excessive Cook's distance values
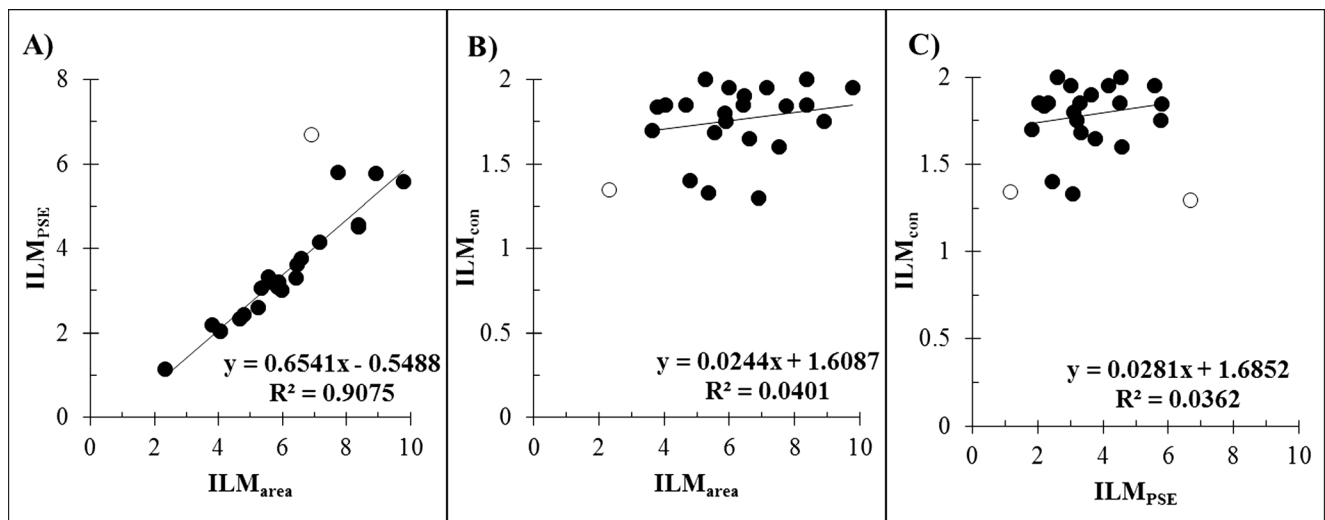
Fig. 5 Correlations between the measures of (A) ILM_{area} and ILM_{pse}, (B) ILM_{area} and ILM_{con}, and (C) ILM_{pse} and ILM_{con} after collapsing over bar onsets/offset conditions and bright/dim cue conditions. Open circles indicate data pairs omitted due to excessive Cook's distance values

explanations need to account for published findings for both onset and offset bars and cannot dismiss a finding based upon this methodological difference alone. The ILM_{con} measure, however, showed poor and nonsignificant test-retest reliability between onset and offset bars, resulting in only weak evidence against the null hypothesis. Examination of the data suggests this reflects ceiling effects limiting the variability as many participants showed measurements near the maximum value of two in at least one of the conditions. It appears that as an index of ILM, ILM_{area} and ILM_{PSE} can be used with more confidence than ILM_{con}. Because ILM_{con} is a common measure in the literature (Christie & Barresi, 2002; Hikosaka et al., 1993c; von Grünau & Faubert,

Fig. 6 Decision times in milliseconds as an exponential decay function of the distance from the PSE for Experiment 1

1994), the poor test-retest reliability of this measure draws into question its suitability as a primary experimental measure.

For example, ILM_{con} did not detect that ILM for onset and offset bars were not equal in magnitude suggesting that saliency had reached ceiling. The ILM_{PSE} measure indicated that ILM for onset bars required slower motion to be cancelled than did offset bars; this was similar to the smaller ILM found for offset bars with the ILM_{area} measure. It should be noted that a difference in illusion magnitude does not necessarily indicate that the illusions for offset bars and onset bars require separate explanations.

The decision time data showed a very strong distance decay function in relation to the PSE. This relationship indicates that as the real and illusory motion signals combine the information upon which the participant makes a decision is reduced, supporting the interpretation that the motion signals are combining (Crawford et al., 2010). The alternative was that the PSE reflects a point of equal winning probability in a system where the illusory signal and the real motion signal compete for priority. However, as both the decision times to pure illusory motion (no real motion but left or right flashes) and to real motion (real left or right motion in the absence of a flash) produce fast response times then the PSE would still reflect fast decisions either to the illusory signal or the real motion signal, whichever happened to win on a given trial. Experiment 2 will further explore the idea of cancelation of motion perception further by including a third response option, where participants can report "no motion." If ILM and real motion signals are competing, rather than concealing, then the no motion option should not be chosen more often as one approaches the PSE.
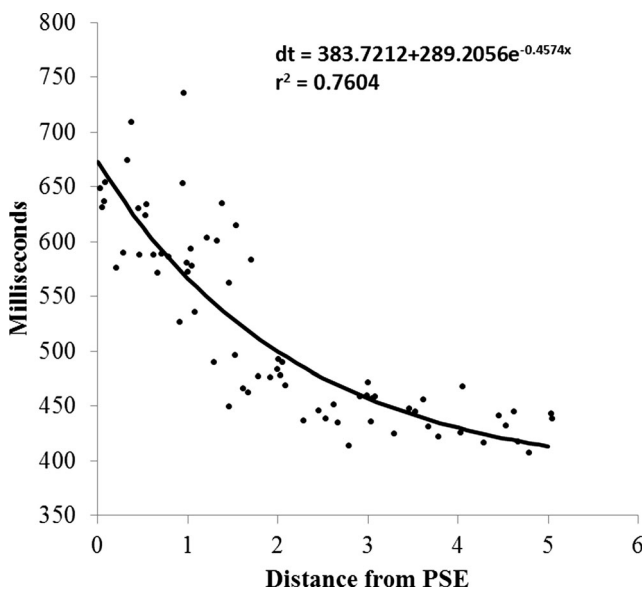
In summary, the results suggest that there is a common illusion being generated in both onset and offset bars following either a brief increase or decrease in the luminance of one of the boxes with the illusory motion being away from this brief luminance fluctuation.

Experiment 2

The purpose of the next experiment was to attempt to replicate and extend the rILM findings from Hamm et al. (2014) to the onset bar condition and to measure rILM using the cancelation procedure. Because rILM is reflected by motion towards the flash, the percept curves should reverse relative to the normal ILM, resulting in negative values for all of the measures. If ILM and rILM reflect a common underlying process then participants who are sensitive to this process and show large positive values during ILM therefore should produce large negative values during rILM displays. Moreover, those with small positive values during ILM should produce small negative values during rILM. In other words, if ILM and rILM both reflect the same underlying process operating in the same way then the prediction is for a strong negative correlation. Alternatively, if ILM and rILM arise from different processes then there is no reason to expect them to be strongly correlated.

An additional goal of Experiment 2 is to further examine if ILM and real motion cancel each other, as is suggested by the decision time data decaying as a function of distance from the PSE. To do this, participants were asked to make both 2AFC responses (left vs. right) and 3AFC (left, right, no motion). The presence of rILM during the flash suggests that the change in the bar segments near the cue are not being detected as quickly as the more distant bar segments. For real motion to cancel rILM, these sections need to be presented before the more distant sections, and so in the rILM condition the bar is removed starting earlier and earlier in time as the real motion speed is reduced. This is so that all conditions complete their removal at the same point during the flash, specifically 1 frame into the cue.

Method

Participants

Twenty-six participants completed the study (17 males; 9 females), but two participants (1 male, 1 female) were excluded due to responses errors (1 inverted the responses for left and right and 1 had chance level performance for real motion only trials). The results are based on the 24 remaining participants (mean age = 21.5 years, SD = 2.11, range = 17–27). Participants were recruited from university undergraduate and postgraduate students. Twenty were right-handed, two

were left-handed, and two were ambidextrous (Oldfield, 1971). The study was approved by the University of Auckland Human Participants Ethics Committee. All participants provided informed, written consent before participation.

Apparatus

Stimulus presentation and data collection was controlled by a desktop Pentium 3 with a 500-MHz processor and an S3 4-MB internal graphics card. The computer had 128-MB RAM and was running Windows 98 and rebooted in MS DOS mode (Myors, 1999). Stimuli were displayed on a 17-inch Phillips Brilliance 17A monitor, running at 60 Hz with a resolution of 640×480 pixels. Responses were made on the number pad of the keyboard. The mean of five luminance readings of the display items taken under the lighting conditions of the experiment using a Konica Minolta LS-100 luminance meter are presented.

A black fixation cross (3.34 cd/m^2) measuring $0.5° \times 0.5°$ was presented in the centre of a neutral grey background (34.48 cd/m^2). The boxes (57.27 cd/m^2) measured $1.9° \times 1.9°$ and were positioned with their centres 1.1^0 above the fixation cross and $4.6°$ to the left and right. When the boxes flashed the luminance increased to 94.06 cd/m^2. The bar (52.27 cd/m^2) measured $7.3° \times 1.4°$ and was centred vertically with respect to the boxes and touched the inner edge of both the left and right box.

Procedure

The procedure was similar to the bright cue and line onset condition of Experiment 1. Participants placed their heads in a chin rest and viewed the stimuli from a distance of 57 cm. A trial began with a 500-ms fixation display, at which point either the left, right, or neither box flashed for 50 ms. As in Experiment 1, upon the offset of the flash the two boxes were joined by the bar with one of the seven levels of real motion, slow left through to slow right, coded as −3 to +3 to reflect both the direction and the number of screen breaks involved in the drawing process. This constitutes the bar after the flash condition. However, in addition to the bar after the flash condition, the bar also could appear such that it was fully presented after the flash had been presented for one of its three screen refreshes, as in the reverse ILM condition of Hamm et al (2014). Although this means in the medium and slow motion conditions that the first segment of the bar was presented before the onset of the flash, this condition is referred to as the bar during the flash condition to reflect the common end point of the bar presentation.

Downing and Treisman (1997) indicate that when the bar precedes the cue this results in rILM. In the no-flash trials, the stimulus conditions resulted in near identical displays and the trials were coded as bar during the flash or after the flash based

on the temporal factors associated with the bar presentation. So, the conditions are not identical and constitute two separate conditions.

Decision timing in all conditions began with the presentation of the first bar segment, regardless of the relationship with the flash. Following a response, there was a 1,000-ms delay before the beginning of the next trial. There were 20 trials in each of the 2 bar conditions (bar after the flash, bar during the flash) x 3 cue locations (left, right, no flash) x 7 levels of motion conditions, for a total of 840 trials in a block. Every participant performed two blocks of trials, one in which only left vs. right motion responses were to be made (on the 4 and 6 key of the number pad for left and right, respectively) and one in which a third response (the 5 key) was included for no motion. The 2AFC and 3AFC blocks were counterbalanced over participants. While percept scores during the 2AFC task were calculated similar to Experiment 1, in the 3AFC block no motion responses were given a score of 0 in addition to the scoring of −1 for left and +1 for right responses.

The experiment was conducted in a well-lit room, one participant at a time. The participants sat on an adjustable chair, with their heads on a chin rest 57 cm from the monitor. Participants were verbally instructed to fixate their eyes on the fixation cross, ignore any flashes, and indicate the direction of any motion in the bar using the 4 and 6 keys on the number pad to indicate left and right, respectively. In the 2AFC task, participants were asked to guess left or right if they were uncertain or perceived no motion and to try to distribute their guesses equally to the left and right keys. For the 3AFC task, participants were able to indicate with and additional key, the 5, if they were uncertain or perceived no motion. The instruction concerning responding in the direction of the initial motion should the display appear to start and then switch directions was not included. The instructions were given verbally and also were available on the screen during the experiment.

Results

Percept scores

Trials in which an invalid key or no key was pressed (1.1 %) or an anticipatory response was made in less than 200 ms (1.2 %) were dropped from the analysis. The percept scores were calculated from the remaining trials. Figure 7 illustrates the percept scores (A-F) and mean decision times (G-L). As with Experiment 1, ILM_{area} was quantified by the area between the left and right flash curves by subtracting the area under the right flash curve from the area under the left flash curve. ILM_{pse} was calculated as the distance between the points of subjective equality for the left and right flash curve by subtracting the PSE for the left flash from the PSE for the right flash condition. ILM_{con} was calculated as the distance

between the observed percept scores in the no-motion condition by subtracting the right flash no-motion percept score from the left flash no-motion percept score. The calculations for all measures were conducted in this direction, because a reverse illusion results in a negative value, which reflects the change in direction.

Quantification of ILM

The ILM_{area}, ILM_{PSE}, and ILM_{con} were analysed in a two-way repeated measures ANOVA, with flash-bar relationship (bar after the flash vs. bar during the flash) and number of responses (2 vs. 3) as factors.

ILM_{area}

There was a main effect of flash-bar relationship ($F_{(1, 23)} = 276.81$, MSE = 2.264, $p < 0.001$, $\eta_p^2 = 0.923$; M = 4.7 vs. −0.41, for bar after the flash vs. bar during the flash, respectively). There was no significant main effect found for the number of response categories ($F_{(1,23)} = 0.22$, MSE = 0.545, $p = 0.64$, $\eta_p^2 = 0.010$; M 2.2 vs. 2.1 for 2AFC vs. 3AFC, respectively). A one sample t test was conducted on the during the flash area measure after combining 2AFC and 3AFC conditions and this did not reach significance ($t_{(23)} = -1.72$, $p = 0.099$), suggesting there was no rILM.

ILM_{PSE}

There was a main effect of flash-bar relationship, ($F_{(1, 23)} = 164.87$, MSE = 1.072, $p < 0.001$, $\eta_p^2 = 0.878$; M = 2.53 vs. −0.19, for bar after the flash vs. bar during the flash, respectively). There was no significant main effect found for the number of response categories ($F_{(1,23)} = 0.28$, MSE = 0.182, $p = 0.60$, $\eta_p^2 = 0.012$; M 1.19 vs. 1.15 for 2AFC vs. 3AFC, respectively). There was no significant interaction between the flash-bar relationship and the number of response categories ($F_{(1,23)} = 1.91$, MSE = 0.138, $p = 0.181$, $\eta_p^2 = 0.077$). A one-sample t test was conducted during the flash ILM_{PSE} measure after combining 2AFC and 3AFC conditions, and this did not reach significance ($t_{(23)} = -1.24$, $p = 0.228$), suggesting there was no rILM.

ILM_{con}

There was a main effect of flash-bar relationship ($F_{(1, 23)} = 315.36$, MSE = 0.171, $p < 0.001$, $\eta_p^2 = 0.932$; M = 1.57 vs. 0.07, for bar after the flash vs. bar during the flash, respectively). There was no significant main effect found for the number of response categories ($F_{(1,23)} = 1.24$, MSE = 0.075, $p > 0.277$, $\eta_p^2 = 0.051$; M = 0.79 vs. 0.85 for 2AFC vs. 3AFC, respectively), and no significant interaction ($F_{(1,23)} = 2.38$, MSE = 0.063, $p = 0.137$, $\eta_p^2 = 0.094$) was found between flash-bar
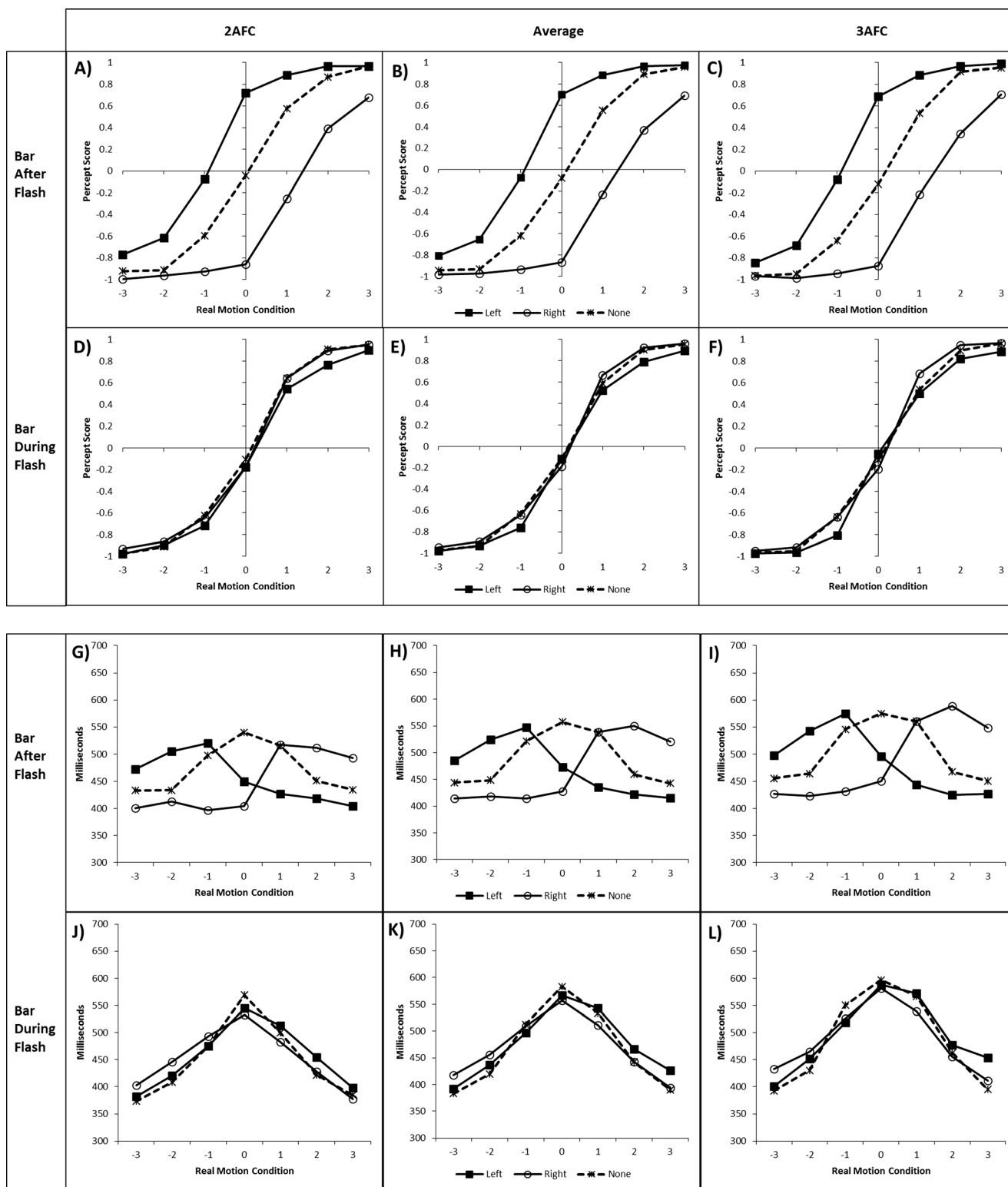
Fig. 7 The top series (A-F) shows the percept scores, and the bottom series (G-L) shows the mean decision times in milliseconds. Within each series, the top row shows the data from the onset bar after the flash condition and the bottom row shows the data from the onset bar during the flash condition. Left-most columns show 2AFC data, right-most column shows the 3AFC data, and the middle column collapses the data between the 2AFC and 3AFC tasks. All figures show data from all 7 real motion speeds ranging from slow left (−3) through to slow right (+3). In the top two rows, positive values on the y-axis indicate rightward motion was perceived more frequently, up to +1 for only perceiving rightward motions, and negative values indicate leftward motion was perceived more frequently up to −1 for only perceiving leftward motions. Solid lines with filled symbols reflect a flash on the left square, solid lines with open circles reflect a flash on the right square, and finally, dashed lines with asterisks reflect the no flash condition

relationship and the number of response categories. A one-sample t test was conducted on the during the flash ILM_{con} measure after combining 2AFC and 3AFC conditions and this did not reach significance ($t_{(23)} = 0.62$, $p = 0.542$), suggesting there was no rILM.

Relationships between the measures

The area, PSE, and consistency measures were averaged over the 2AFC and 3AFC designs to examine relationships between the measures. Because there are three pairings for the bar after the flash and bar during the flash condition (Fig. 8), significance is determined at $p < 0.0167$. After removing data pairs with excessive Cook's distance the following relationships were found: when the bar was presented after the flash, ILM_{area} was significantly correlated with ILM_{PSE} (Fig. 8A: $r_{(21)} = 0.97$, $p < 0.001$, $p_{H0|D} < 0.001$) and with ILM_{con} (Fig. 8B: $r_{(22)} = 0.68$, $p = 0.002$, $p_{H0|D} = 0.028$), which is considered very strong and strong evidence against the null hypothesis, respectively. In addition, ILM_{PSE} was significantly correlated with ILM_{con} (Fig. 8C: $r_{(21)} = 0.59$, $p < 0.001$, $p_{H0|D} = 0.004$) and showed very strong evidence against the null hypothesis. If the outlier data pairs are included, the corresponding r values are 0.97, 0.59, and 0.52. Even in the absence of illusory motion in the bar during the flash condition the measures are correlated. ILM_{area} was correlated with ILM_{PSE} (Fig. 8D: $r_{(20)} = 0.99$, $p < 0.001$, $p_{H0|D} < 0.001$), ILM_{area} was correlated with ILM_{con} (Fig. 8E: $r_{(20)} = 0.86$, $p < 0.001$, $p_{H0|D} < 0.001$), and ILM_{PSE} was correlated with ILM_{con} (Fig. 8F: $r_{(20)} = 0.89$, $p < 0.001$, $p_{H0|D} < 0.001$), all of which are considered very strong evidence against the null hypothesis. If the outlier data pairs are included, the corresponding r values are 0.98, 0.78, and 0.74. Using Fischer's z-transformation (Cohen & Cohen, 1983) to compare the strength of the relationship between the bar after the flash and bar during the flash conditions and a Bonferroni corrected significance criterion of 0.0167, suggested the strength of the correlations did not differ when ILM_{area} predicted ILM_{PSE} ($z = 1.45$, $p = 0.07$) or when ILM_{area} predicted ILM_{con} ($z = 1.51$, $p = 0.06$) or when ILM_{PSE} predicted ILM_{con} ($z = 2.297$, $p = 0.018$).

Decision time

The mean decision times after removing anticipation trials with decision times less than 200 ms (1.16 %) and trials in which either no response was made during the trial or an invalid response key was pressed (1.14 %) are shown in the lower half of Fig. 7G-L. It should be noted that as the presentation of the line

Fig. 8 Upper row depicts the correlations during the onset bar after the flash conditions between (A) ILM_{area} and ILM_{PSE}, (B) ILM_{area} and ILM_{con}, and (C) ILM_{PSE} and ILM_{con}. Lower row depicts the correlations during the onset bar during the flash conditions between (D) ILM_{area} and ILM_{PSE}, (E) ILM_{area} and ILM_{con}, and (F) ILM_{PSE} and ILM_{con}. Open circles indicate data pairs omitted due to excessive Cook's distance values

occurs at different points of time following the start of the trial, the no-flash conditions are not identical with respect to the during-the-flash and after-the-flash conditions and are not combined.

As with Experiment 1, there was an exponential decay in the decision times as the condition moved further from the PSE (Fig. 9). The best fitting equation was dt $= 348.5181 + 233.9284e^{-0.4004x}$, $r_{(81)} = 0.8777$, $p < 0.05$, $p_{H0|D} < 0.001$, which is very strong evidence against the null hypothesis. As in Experiment 1, the exploratory congruency effect was examined and can be found in Appendix B.

Peakedness of the decision times

After combining the no flash data from the bar after the flash and the bar during the flash conditions (Fig. 10A), peakedness values for the 2AFC and 3AFC blocks were calculated from the decision times by subtracting the mean decision time to fast left and fast right real motion trials from the decision time during no real motion trials. These were then compared by a within-subjects t test. The 2AFC condition was found to produce a more peaked function ($t_{(23)} = 2.35$, $p = 0.028$), with the mean differences shown in Fig. 10B.

Use of the third response option (no motion)

Use of the no-motion response varied between the participants, ranging between 0 and 215 (0.0 % to 25.6 % of the 3AFC trials) with the mean number of no-motion responses for each of the bar, cue, and motion conditions shown in Fig. 11A and B. The number of times the no motion response was made during the 3AFC task was analysed in a three-way repeated measures ANOVA with bar condition (2; during vs. after the flash), cue (3; left, right, none), and motion (7; slow left through to slow right) as factors. There was no main effect of bar condition ($F_{(1,23)} = 2.21$, MSE $= 2.665$, $p = 0.15$, $n_p^2 = 0.088$). There was a main effect of cue ($F_{(1,23)} = 41.28$, MSE $= 7.713$, $p < 0.001$, $n_p^2 = 0.642$). There was a main effect of motion ($F_{(6,138)} = 43.21$, MSE $= 14.799$, $p < 0.001$, $n_p^2 = 0.653$).

The interaction between bar and cue did not reach significance ($F_{(2,46)} = 2.53$, MSE $= 2.698$, $p = 0.090$, $n_p^2 = 0.099$). The interaction between bar and motion

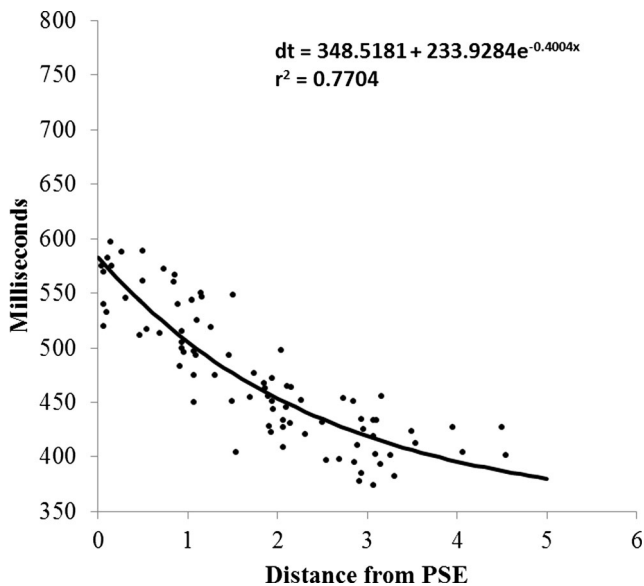$$dt = 348.5181 + 233.9284e^{-0.4004x}$$
$$r^2 = 0.7704$$
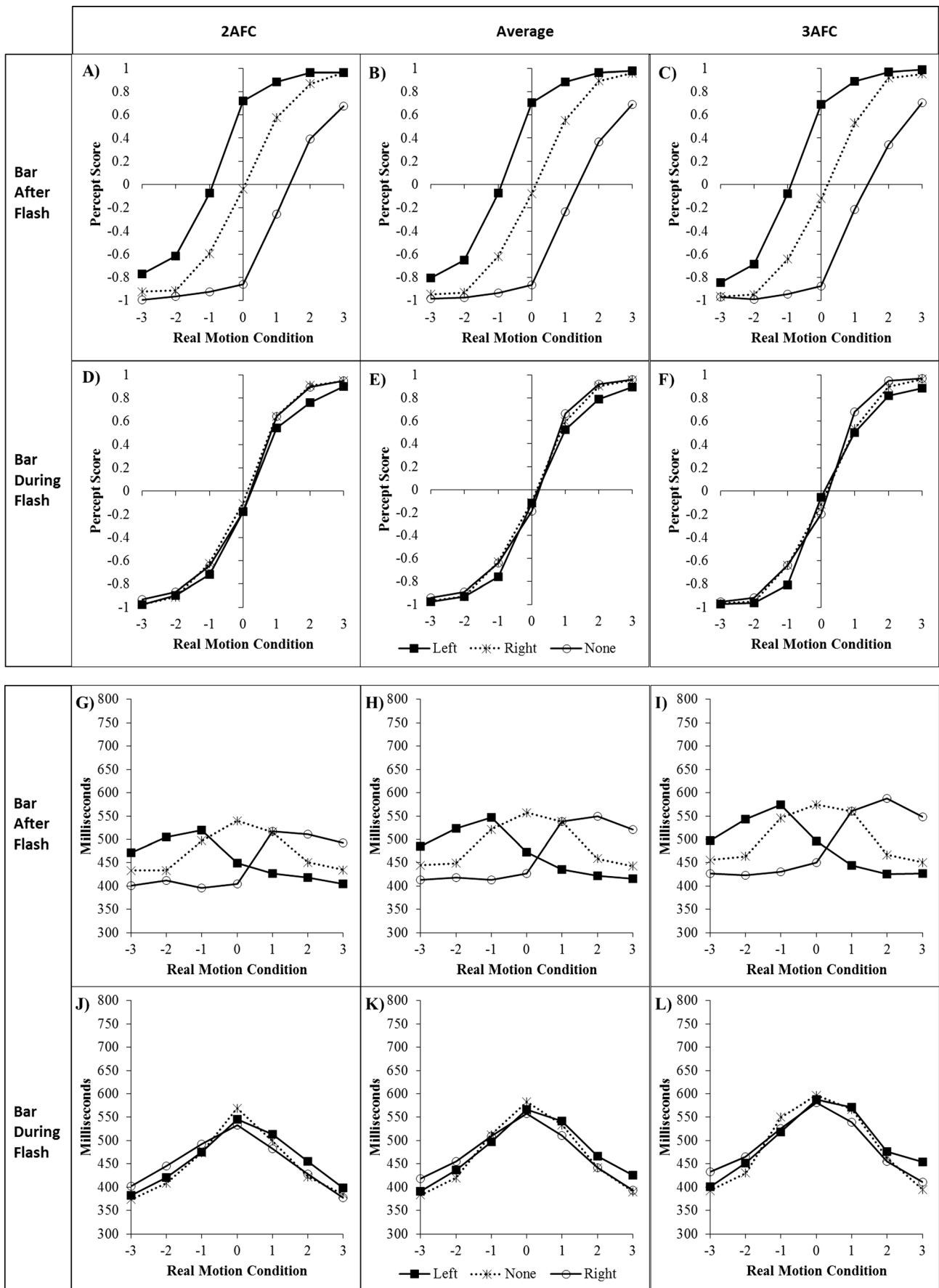
Fig. 9 Decision times in milliseconds as an exponential decay function of the distance from the PSE for Experiment 2

was significant ($F_{(6,138)}$ = 6.14, MSE = 2.705, $p <$ 0.001, n_p^2 = 0.211). The interaction between cue and motion was significant ($F_{(12,276)}$ = 37.15, MSE = 4.919, $p < 0.001$, n_p^2 = 0.618). Finally, the three way interaction was significant ($F_{(12,276)}$ = 23.56, MSE = 2.575, $p < 0.001$, n_p^2 = 0.506).

Discussion

Neither the 2AFC nor the 3AFC condition replicated the rILM effect with onset bars (see percept scores in Fig. 7A-F). Although both ILM$_{area}$ and ILM$_{pse}$ showed a negative overall mean, which is expected if rILM occurs, neither condition was

significantly different from 0, indicating the failure to conform to the prediction that rILM will extend to onset bars. Moreover, ILM$_{con}$ was positive overall suggesting ILM rather than rILM, but this too did not significantly differ from zero. These findings against illusory motion during the flash are the results Hamm et al. (2014) originally anticipated, because this was to be their control condition. They argued that a bar change so soon after the flash onset would occur before there was time to establish an attentional gradient. Given that Experiment 1 established that ILM for onset and offset bars are highly related, the failure for onset bars to produce rILM suggests that rILM does not reflect the same processes as ILM.

In regard to the secondary research aim, there appears to be no difference in the percept scores based upon using a 2AFC or 3AFC procedure (Fig. 7). However, the overall decision times were slower in the 3AFC task than in the 2AFC task, indicating that the decision times can be influenced by factors that do not influence perceptual measures of ILM. In addition, the decision time functions during the no flash condition were more peaked in the 2AFC task than in the 3AFC task, which is consistent with the 2AFC task involving a single response boundary, whereas the 3AFC task is likely to have two response boundaries—one between left motion and no motion and another between no motion and rightward motion (Cartwright, 1941).

Before considering the findings of Experiment 2 any further, it was important to determine if the lack of rILM was due to the use of onset bars or if rILM does not occur in the context of a cancellation paradigm. A failure to replicate rILM with offset bars would suggest that the finding in Hamm et al. (2014) is unstable. Therefore, for experiment three we replicated the procedures in Experiment 2 except the display began
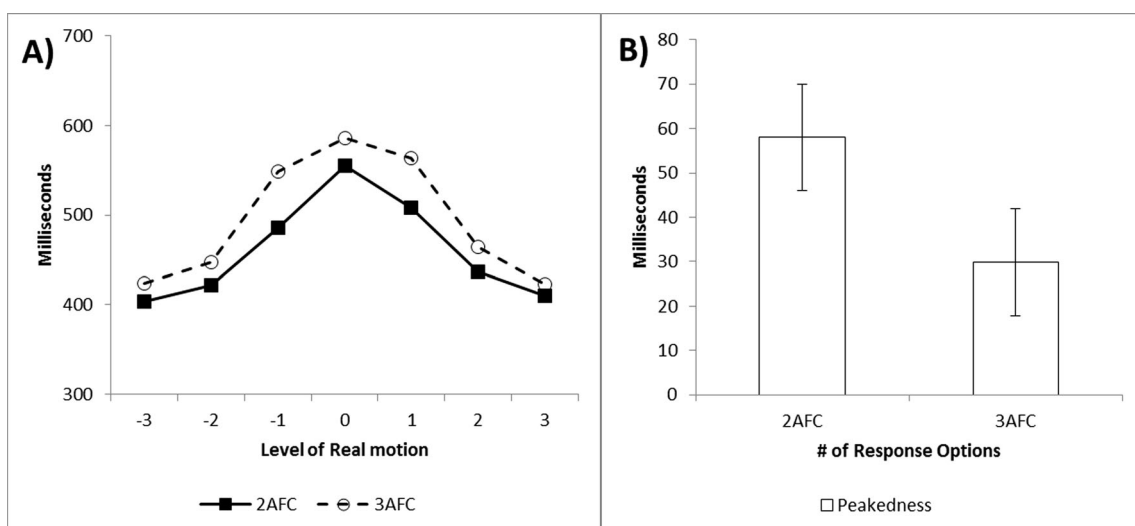

Fig. 10 Mean response time for the no flash conditions with onset bars during the 2AFC and 3AFC conditions (A). Peakedness as indexed by the difference between the averages of the fast left and right motion

conditions and the no motion response times for the 2AFC and 3AFC is shown on the right (B). Error bars show the standard error of the difference as a within subjects error bar (Masson & Loftus, 2003)
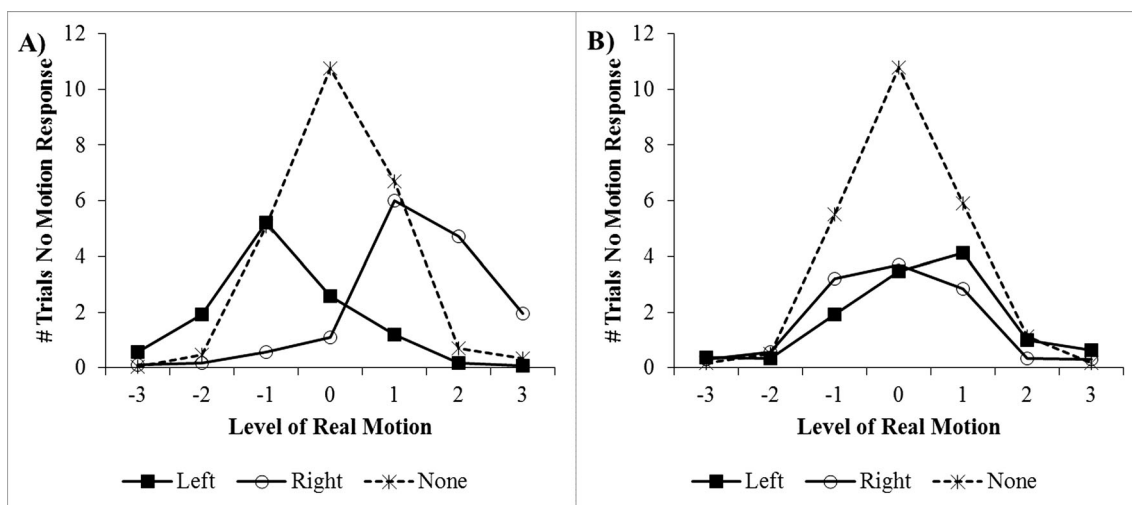
Fig. 11 Mean number of no motion responses for left (solid lines with filled squares), right (solid lines with open circles), and no flash conditions (dotted lines with asterisks) at each of the 7 levels of real motion for (A) the onset bar after the flash and (B) the onset bar during the flash conditions

with the bar present and the motion was created by sequential removal of the bar segments.

Experiment 3

Methods

Participants

Twenty-four participants completed the study (13 males; 11 females; mean age 21.74 years). Seven participants (6 males, 1 female) had participated in Experiment 2 with a minimum of 86 days between experiments (M = 185.9 days). The remainder had not participated in either of the previous studies. Participants were recruited from university undergraduate and postgraduate students. Twenty-two were right-handed and two were left-handed (Oldfield, 1971). The study was approved by the University of Auckland Human Participants Ethics Committee. All participants provided informed, written consent before participation.

Stimuli and apparatus

The same stimuli and apparatus were employed as in Experiment 2 with only the following change. The bar was presented at the same time as the two boxes and was removed in sections to give the appearance of a bar disappearing, similar to a train going into a tunnel rather than emerging from it.

Procedure

The same procedure was followed as with Experiment 2.

Results

Percept scores

Trials in which a response was made in less than 200 ms (1.2 %) were deemed to be anticipations and were dropped from the analysis, as were trials in which either an invalid or no response was made (1.2 %). The percept scores were calculated from the remaining trials (97.6 %). Figure 12 illustrates the percept scores (A-F) and mean decision times (G-L).

Quantification of ILM

As in Experiment 2, ILM_{area}, ILM_{pse}, and ILM_{con} were analysed in a two-way repeated measures ANOVA, with flash-bar relationship (bar after the flash vs. bar during the flash) and number of responses (2 vs. 3) as factors. Again, negative values for these measures indicate the illusion was towards the flash and positive values indicate an illusion away from the flash.

ILM_{area}

There was a main effect of flash-bar relationship ($F_{(1, 23)}$ = 135.61, MSE = 6.740, $p < 0.001$, $\eta_p^2 = 0.8.55$; M = 3.83 vs. −2.33, for bar after the flash vs. bar during the flash, respectively). There was no significant main effect found for the number of response categories ($F_{(1,23)}$ = 0.54, MSE = 0.218, $p = 0.47$, $\eta_p^2 = 0.023$; M 0.72 vs. 0.78 for 2AFC vs. 3AFC, respectively). There was no significant interaction between the flash-bar relationship and the number of response categories ($F_{(1,23)}$ = 3.02, MSE = 0.600, $p = 0.10$, $\eta_p^2 = 0.116$).

ILM_{pse}

There was a main effect of flash-bar relationship ($F_{(1, 23)}$ = 56.77, MSE = 4.916, $p < 0.001$, $\eta_p^2 = 0.712$; M = 2.23 vs. −1.18, for bar after the flash vs. bar during the flash, respectively). There was no significant main effect found for the number of response categories ($F_{(1,23)}$ = 0.16, MSE = 0.159, $p = 0.69$, $\eta_p^2 = 0.007$; M 0.54 vs. 0.51 for 2AFC vs. 3AFC, respectively). There was no significant interaction between the flash-bar relationship and the number of response categories ($F_{(1,23)}$ = 3.58, MSE = 0.380, $p = 0.07$, $\eta_p^2 = 0.135$).

ILM_{con}

There was a main effect of flash-bar relationship, ($F_{(1, 23)}$ = 511.63, MSE = 0.281, $p < 0.001$, $\eta_p^2 = 0.957$; M = 1.48 vs. −0.97, for bar after the flash vs. bar during the flash, respectively). There was no significant main effect found for the number of response categories ($F_{(1,23)}$ = 0.85, MSE = 0.076, $p = 0.367$, $\eta_p^2 = 0.036$; M = 0.23 vs. 0.28 for 2AFC vs. 3AFC, respectively), and no significant interaction ($F_{(1,23)}$ = 1.28, MSE = 0.109, $p = 0.270$, $\eta_p^2 = 0.053$) was found between flash-bar relationship and the number of response categories.

Relationships between the measures

The area, PSE, and consistency measures were averaged over the 2AFC and 3AFC designs to examine relationships between the measures. As with Experiment 2, significance was determined at $p < 0.0167$ to correct for the three comparisons made within each line condition. Data pairs with extreme Cook's distance values were removed from consideration. The scatter plots may be seen in Fig. 13. When the bar was removed after the flash, ILM_{area} was significantly correlated with ILM_{PSE} (Fig. 13A: $r_{(20)}$ = 0.97, $p < 0.001$, $p_{H0|D} < 0.001$, very strong evidence against the null hypothesis) but neither ILM_{area} (Fig. 13B: $r_{(21)}$ = 0.33, $p = 0.128$, $p_{H0|D} = 0.5602$, weak evidence in support of the null hypothesis) nor ILM_{PSE} (Fig. 13C: $r_{(21)}$ = 0.14, $p = 0.510$, $p_{H0|D} = 0.7925$, positive evidence in favour of the null hypothesis) was significantly correlated with ILM_{con}. When the bar was removed during the flash, ILM_{area} was correlated with ILM_{PSE} (Fig. 13D: $r_{(18)}$ = 0.99, $p < 0.001$, $p_{H0|D} < 0.001$) and both ILM_{area} (Fig. 13E: $r_{(20)}$ = 0.82, $p < 0.001$, $p_{H0|D} < 0.001$), and ILM_{PSE} (Fig. 13F: $r_{(20)}$ = 0.79, $p < 0.001$, $p_{H0|D} < 0.001$) was correlated with ILM_{con}, all of which produced very strong evidence against the null hypothesis. Comparison of the relationships between the bar after the flash and bar during the flash conditions suggested that the relationship between ILM_{area} and ILM_{PSE} did not differ between conditions (z = 1.49, $p = 0.07$), but

Within each series, the top row shows the data from the offset bar after the flash condition and the bottom row shows the data from the offset bar during the flash condition. Left-most columns show 2AFC data, right-most column shows the 3AFC data, and the middle column collapses the data between the 2AFC and 3AFC tasks. All figures show data from all 7 real motion speeds ranging from slow left (−3) through to slow right (+3). In the top two rows, positive values on the y-axis indicate rightward motion was perceived more frequently, up to +1 for only perceiving rightward motions, and negative values indicate leftward motion was perceived more frequently up to −1 for only perceiving leftward motions. Solid lines with filled symbols reflect a flash on the left square, solid lines with open circles reflect a flash on the right square, and finally, dashed lines with asterisks reflect the no flash condition

the relationship with ILM_{con} was stronger (Bonferroni corrected criterion of $p < 0.0167$) in the bar during the flash condition for both ILM_{area} (z = 2.59, $p = 0.005$) and ILM_{PSE} (z = 2.90, $p = 0.002$). If the outlier data pairs are included, then the values become r = 0.94, 0.45, and 0.26 for the bar after the flash condition, and r = 0.96, 0.44, and 0.32 for the bar during the flash condition.

Comparison of ILM and rILM

Data were collapsed over the number of responses, and the ILM_{area}, ILM_{PSE}, and ILM_{con} for each individual from the bar after the flash was correlated with their corresponding value from the bar during the flash condition to determine if there was any evidence to suggest that ILM and rILM were mediated by a common underlying process (Fig. 14). Because each of these comparisons effectively asks the same question, a Bonferroni correction factor of 3 is applied, resulting in significance being determined at $p < 0.0167$. After removal of the outlier data pairs based on Cook's distance, none of the relationships reached significance; ILM_{area} (Fig. 14A: $r_{(20)}$ = −0.18, $p = 0.423$, $p_{H0|D}$ = 0.7655, positive evidence in favour of the null hypothesis), ILM_{PSE} (Fig. 14B: $r_{(20)}$ = −0.43, $p = 0.046$, $p_{H0|D}$ = 0.3311, weak evidence against the null hypothesis), and ILM_{con} (Fig. 14C: $r_{(21)}$ = 0.16, $p = 0.466$, $p_{H0|D}$ = 0.7807, positive evidence in favour of the null hypothesis). If the extreme pairs are included, the values become ILM_{area} (r = −0.44), ILM_{PSE} (r = −0.64), and ILM_{con} (r = 0.24).

Decision time

The mean decision times can be seen in the lower half of Fig. 12G-L. As with Experiments 1 and 2, the decision times showed an exponential decay as a function of the distance from the PSE (Fig. 15). The data were best described by the equation dt = 346.3505 + 220.1266$e^{-0.5559x}$, $r_{(81)}$ = 0.8869, $p < 0.05$, $p_{H0|D} < 0.001$, which is very strong evidence against the null hypothesis. Similar to
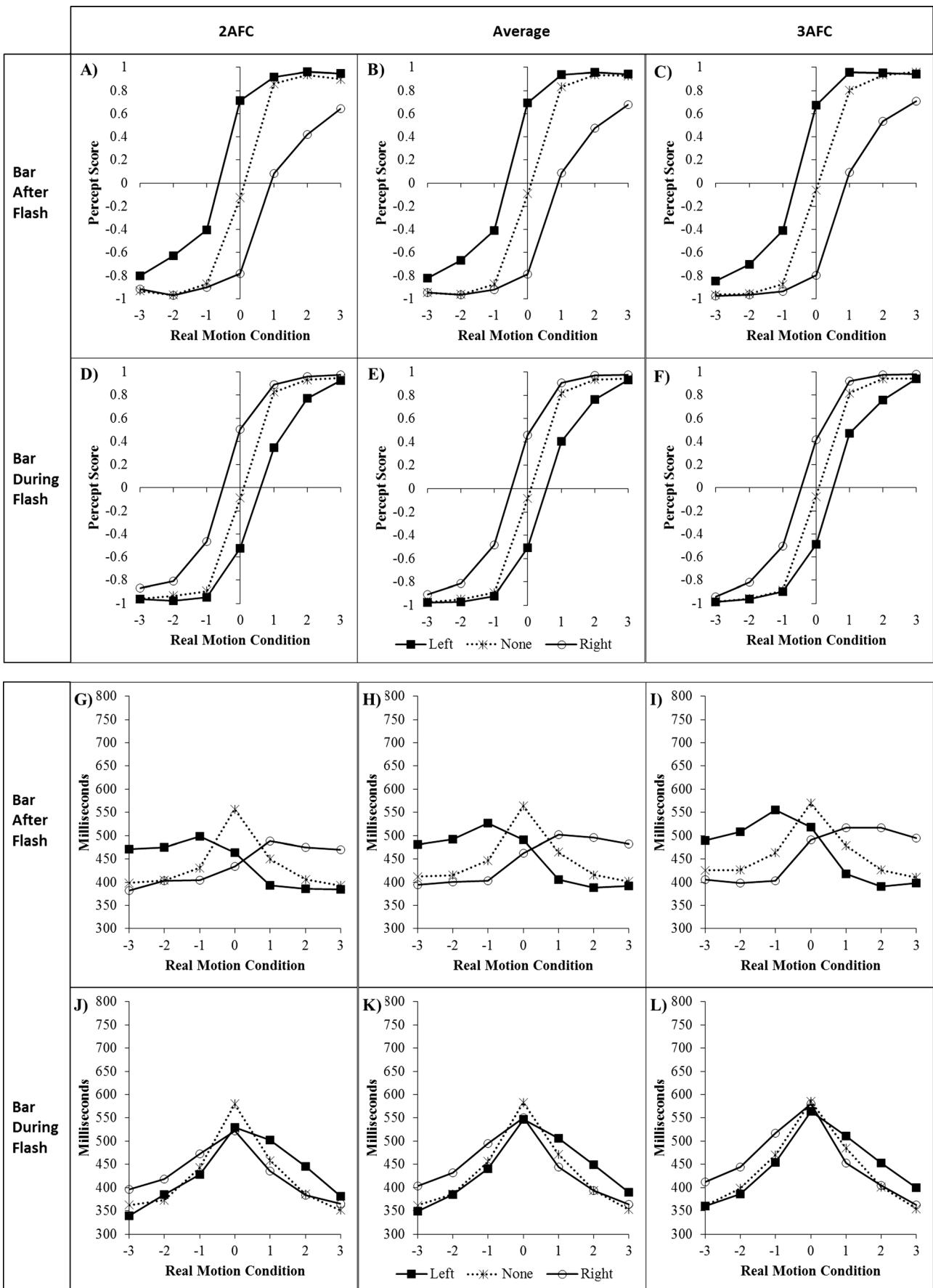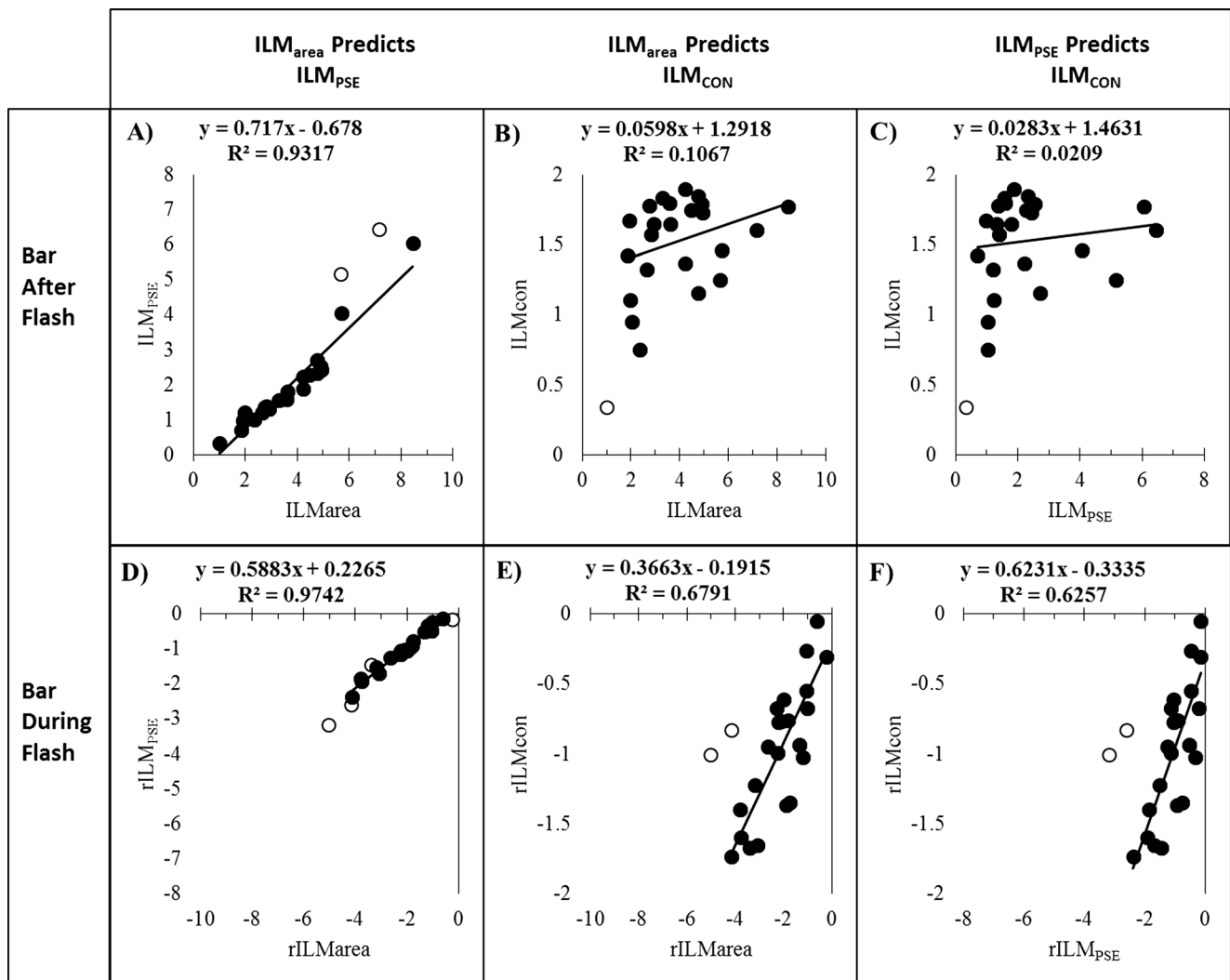
Fig. 13 Upper row depicts the correlations during the offset bar after the flash conditions between (A) ILM_{area} and ILM_{PSE}, (B) ILM_{area} and ILM_{con}, and (C) ILM_{PSE} and ILM_{con}. Lower row depicts the correlations during the offset bar during the flash conditions between (D) ILM_{area} and ILM_{PSE}, (E) ILM_{area} and ILM_{con}, and (F) ILM_{PSE} and ILM_{con}. Open circles indicate data pairs omitted due to excessive Cook's distance values

Experiments 1 and 2, the analysis of the congruency effect may be found in Appendix C.

Peakedness of the decision times

After combining the no flash data from the bar after the flash and the bar during the flash conditions (Fig. 16A), peakedness values for the 2AFC and 3AFC blocks were calculated from the decision times for the no flash trials by subtracting the mean of the fast left and fast right real motion decision time from the no real motion decision time. These were then compared by a within-subjects t test. Although the 2AFC condition tended to produce a more peaked function as was found in Experiment 2, this did not reach significance ($t_{(23)} = 1.73$, $p = 0.098$), with the mean differences shown in Fig. 16B.

Use of the third response option (no motion)

Use of the no-motion response varied between the participants, ranging between 0 and 193 (0.0 % to 23.0 % of the 3AFC trials) with the mean number of no motion responses for each of the bar, cue, and motion conditions shown in Fig. 17A and B. The number of times the no-motion response was made during the 3AFC task was analysed in a three-way repeated measures ANOVA with bar condition (2; during vs. after the flash), cue (3; left, right, none), and motion (7; slow left through to slow right) as factors. There was a main effect of bar condition ($F_{(1,23)} = 4.74$, MSE = 3.758, $p = 0.040$, $n_p^2 = 0.171$). There was a main effect of cue ($F_{(1,23)} = 22.28$, MSE = 3.809, $p < 0.001$, $n_p^2 = 0.492$). There was a main effect of motion ($F_{(6,138)} = 30.31$, MSE = 16.374, $p < 0.001$, $n_p^2 = 0.569$).
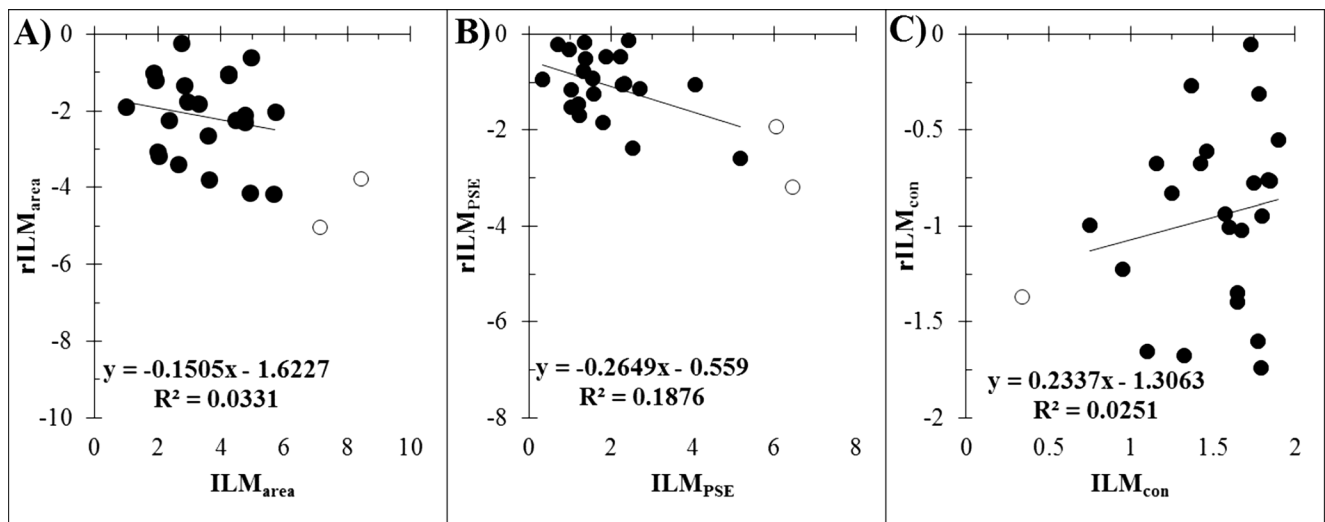
Fig. 14 Correlation between (A) ILM$_{area}$ and rILM$_{area}$, (B) ILM$_{PSE}$ and r ILM$_{PSE}$, and (C) ILM$_{con}$ and rILM$_{con}$ for offset bars after the flash and offset bars during the flash. Open circles indicate data pairs omitted due to excessive Cook's distance values

The interaction between bar and cue did not reach significance ($F_{(2,46)} = 1.14$, MSE = 3.010, $p = 0.329$, $n_p^2 = 0.047$) nor did the interaction between bar and motion ($F_{(6,138)} = 1.06$, MSE = 1.665, $p = 0.389$, $n_p^2 = 0.044$). The interaction between cue and motion was significant ($F_{(12,276)} = 26.13$, MSE = 4.545, $p < 0.001$, $n_p^2 = 0.532$). Finally, the three-way interaction was significant ($F_{(12,276)} = 8.16$, MSE = 3.174, $p < 0.001$, $n_p^2 = 0.262$).

Discussion

Experiment 3 replicated the procedure of Experiment 2 except ILM was generated via offset bars rather onset bars. When the bar was removed after the flash robust ILM away from the

Fig. 15 Decision times in milliseconds as an exponential decay function of the distance from the PSE for Experiment 3

flash was found, replicating the ILM findings that resulted for both onset and offset bars in Experiment 1 and with onset bars in Experiment 2. However, Experiment 3 also replicated the rILM finding of Hamm et al (2014) with motion towards the flash when the bar was removed during the flash, a result that did not occur with onset bars in Experiment 2. In addition, as with Experiment 2, the 2AFC and 3AFC paradigms produce similar results with respect to the percept measures of ILM and rILM. The increased complexity of the 3AFC decision space tended to influence the decision time data in ways consistent with expectations based on assuming there are multiple decision boundaries (Cartwright, 1941), although the peakedness of the 2AFC and 3AFC response time functions did not significantly differ in Experiment 3.

The ILM measures from the 2AFC and 3AFC were strongly correlated. Similarly, the rILM measures from the 2AFC and 3AFC conditions were correlated. This indicates that the additional complexity of the response options did not change the decision space to alter the estimate of the motion percept. However, the measures for ILM and rILM were not correlated with each other, suggesting that the illusion generated after the flash was unrelated to the illusion generated during the flash. This was further tested by comparing the relationship between onset and offset bars from Experiment 1 with the relationship between ILM and rILM from Experiment 3. This analysis indicated that the relationship was stronger in Experiment 1 than Experiment 3 for both ILM$_{area}$ ($z = 2.92$, $p = 0.004$) as well as for ILM$_{PSE}$ ($z = 2.62$, $p = 0.008$) but not for ILM$_{con}$ ($z = 0.89$, $p = 0.373$). While the ILM$_{con}$ for onset and offset bars were not correlated in Experiment 1, many of the data points were at ceiling performance levels, which made interpretation difficult. In Experiment 3, there was less evidence of a ceiling effect, and yet the ILM$_{con}$ and rILM$_{con}$ were still not correlated.
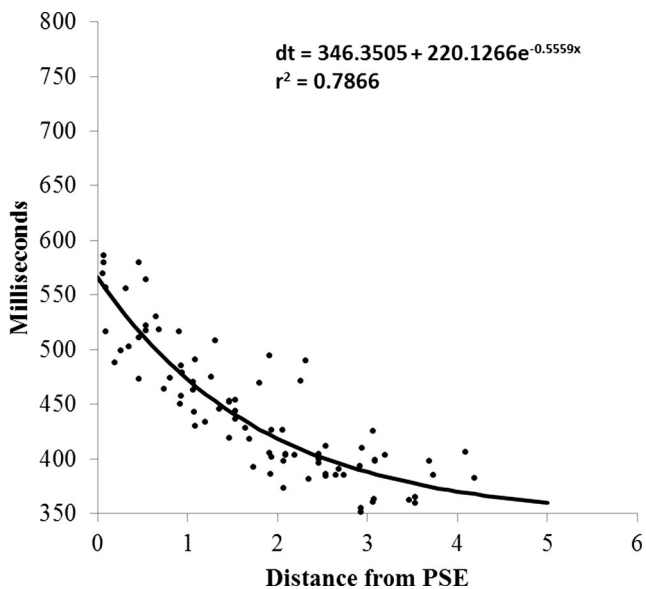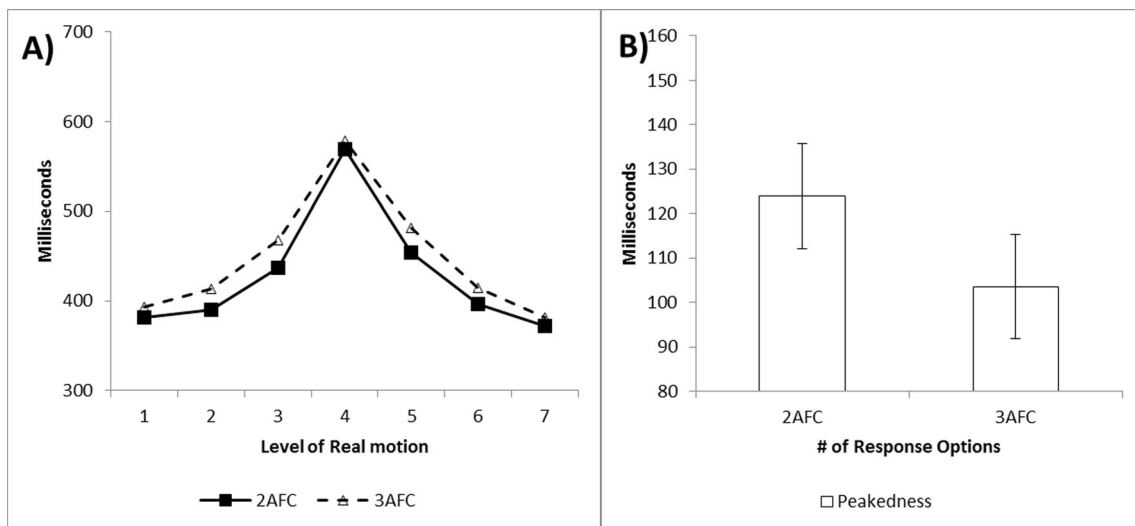
Fig. 16 Mean response time for the no flash conditions with offset bars during the 2AFC and 3AFC conditions (A). Peakedness as indexed by the difference between the averages of the fast left and right motion conditions and the no-motion response times for the 2AFC and 3AFC is shown on the right (B). Error bars show the standard error of the difference as a within subjects error bar (Masson & Loftus, 2003)

In summary, the findings from Experiment 3 suggest that ILM and rILM are not the same illusion and they do not reflect a common underlying process in the way that the ILM data from onset and offset bars do. Rather, the data suggest that two separate explanations are required: one for motion away from the flash in both onset and offset bars that occurs when the bar change occurs after the flash, and one for motion towards the flash in offset, but not onset, bars when the bar change completes during the flash.

As a final interesting but cautious supplementary analysis, it should be noted that Experiment 1 demonstrated that there was a strong positive relationship between the illusions for onset and offset bars as measured by the area between the curves for left and right flash conditions. Seven subjects

participated in both Experiments 2 and 3 and therefore provided ILM data for both onset and offset bars after a substantial amount of time between the two experiments (minimum was 80+ days). While this is an admittedly small sample size, given the strength of the relationship found in Experiment 1, it is of interest to explore the possibility that this individual difference measure for the illusion is stable over long periods of time. Figure 18 shows the relationship between the area measures for onset (Experiment 2) and offset (Experiment 3) ILM from the after-the-flash and, for comparison, rILM$_{area}$ from the during-the-flash conditions. The relationship is strong, and in the direction expected, in the bar after-the-flash condition when both experiments produced robust ILM ($r_{(5)} = 0.74$, $p = 0.057$, $p_{H0|D} = 0.1415$, positive evidence
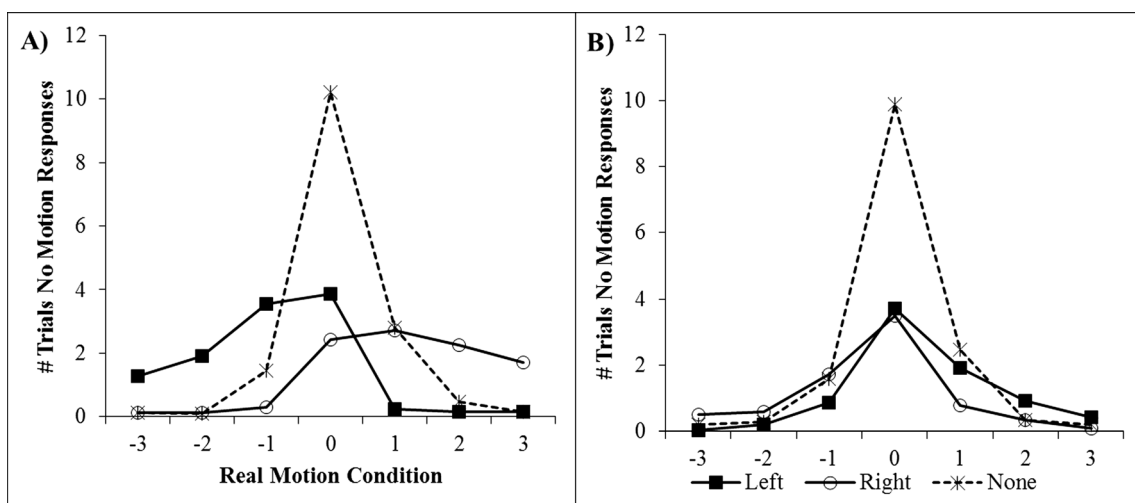

Fig. 17 Mean number of no motion responses for left (solid lines with filled squares), right (solid lines with open circles), and no flash conditions (dotted lines with asterisks) at each of the 7 levels of real motion for (A) the offset bar after the flash and (B) the offset bar during the flash conditions
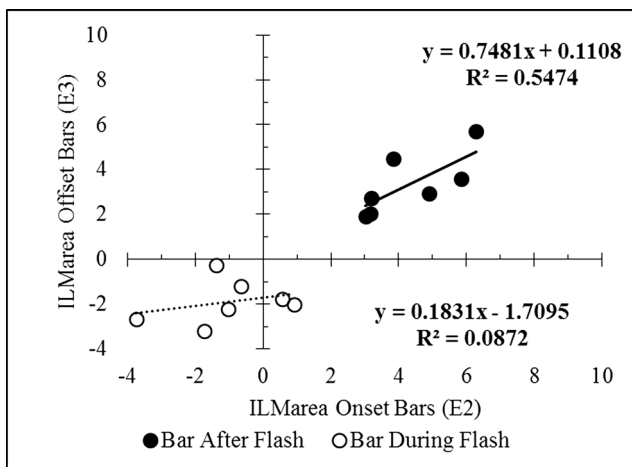
Fig. 18 Correlations between ILM_{area} from the seven who participated in both the onset (E2) and offset (E3) bar experiments when the bar was presented after the flash (solid lines with filled squares) and during the flash (dashed line with open circles)

against the null hypothesis). However, there was no suggestion of a relationship in the measures from the bar during-the-flash conditions, which produced no evidence for rILM in Experiment 2 but strong rILM in Experiment 3 ($r_{(5)} = 0.30$, $p = 0.520$, $p_{H0|D} = 0.6554$, weak evidence in favour of the null hypothesis). While this finding is suggestive, it should be noted that comparison between these correlations using Fischer's z-transformation (Cohen & Cohen, 1983) does not indicate that they are significantly different from each other ($z = 0.9133$, $p = 0.361$) and so the pattern of significance and support for and against the null hypothesis shown above should not be taken as evidence of a significant interaction. It is acknowledged that this analysis is greatly underpowered, and these findings are offered only as an encouraging basis for a larger study designed to examine the question of how stable are these individual measures over extended periods of time. The current data encourage the contemplation that when ILM occurs these are quite stable individual characteristics, at least in adults.

General discussion

Quantifying ILM

The results presented demonstrate the usefulness of examining the similarity of ILM under different conditions at the individual participant level. The illusion was quantified by the area between the curves from the left and right flash conditions, the distance between the points of subjective equality, and the distance between the percept scores for the left and right flash conditions. Whereas the last of these measures, ILM_{con}, is prone to ceiling effects, the former two measures do not appear to be and moreover were highly correlated with

each other. When the consistency measure was not at ceiling, it tended to be correlated with the other two measures, although generally more so for the bar during the flash conditions (no ILM and rILM conditions) than bar after the flash conditions (ILM conditions). ILM_{PSE} and ILM_{con} theoretically quantify orthogonally separate concepts, namely motion speed and motion saliency, whereas the area measure captures a combination of these two concepts. Despite that, the area measure is generally more strongly correlated to ILM_{PSE} than ILM_{con}, again apparently due to ILM_{con} tending to reach ceiling.

In general, most questions about what affects ILM or rILM can be addressed by only examining one of these quantities. This is beneficial, because if one examines all three quantities, then one should report all three analyses and apply a Bonferroni correction for the multiple comparisons. These are multiple comparisons, because each of these tests is effectively asking the same question: Does a given manipulation change ILM?

Unless there are specific theoretical reasons to choose otherwise, it is recommended that the measure ILM_{area} be preferred for the following reasons. First, ILM_{area} will be sensitive to changes in both the speed and saliency dimension and so will be prone to detect the influence of a manipulation regardless of which dimension it primarily influences. Second, while ILM_{PSE} and ILM_{con} are theoretically orthogonal measures, in practice they generally change together until ILM_{con} reaches ceiling. Third, ILM_{area} is derived directly from all of the observed data while ILM_{con} is based only on a subset of the data and ILM_{PSE} requires the fitting of a function, which adds another layer of complications between the interpretation and the observed phenomenon.

The fitting of a function to the percept scores in order to extract ILM_{PSE} opens the door for many additional experimenter degrees of freedom, such as what function to fit (log linear as currently employed or Weibull as in von Grünau, Racette et al., 1996), how to fit it (least squares, maximum likelihood), and the criterion for deciding if a function provides a good fit. All of these decisions open up additional questions of appropriateness of the choice, which can complicate interpretation, become a source of distraction, and reduce the simplicity of comparison between studies. However, if one chooses to fit a psychometric function to individual participants' data it is recommended that whatever the function the fit should be based on least squares as the goal of fitting a function in this situation is to obtain a description of the observed data (Cohen & Cohen, 1983) that will then be subjected to inferential analyses rather than making inferences about the distribution of the underlying function. That being said, we see no obvious reason to expect that the choice of function or method of individual curve fitting would impact greatly upon the results, but this is an area that is worthy of focused examinations.

With respect to the cancelation procedure, a simple 2AFC between left and right motion seems preferable to the 3AFC version that includes the no motion option as well. The percept scores produce near identical functions and the third response option only increases the complexity of the decision space by creating two response boundaries: one between left and no motion and another between no motion and right (Cartwright, 1941). Granted the use of the no motion response directly queries the participants to indicate if the real motion and the ILM cancel each other when they are in opposition, but this same conclusion can also be drawn based on the decision times from the 2AFC protocol (Crawford et al., 2010). Finally, use of the no motion response varied greatly between participants with some never using it at all, even in the no motion no flash condition. This may reflect individual differences with respect to decision biases or differences in motion perception. While the underlying reasons for these differences are interesting questions for research, they are not of primary concerns to studies interested in examining the processes underling ILM and therefore the simpler decision space of the 2AFC procedure is recommended.

Decision times conform to a distance decay function away from the PSE, consistent with an increasingly stronger motion signal. Focused investigations into the distance decay functions obtained from this task should uncover additional information that could help to understand ILM and rILM more completely. For example, it remains to be determined if the parameters of the decay functions are stable at an individual level. If so, these parameters may index underlying psychological constructs that are interest for targeted research.

Considering the individual differences approach

The main goal of the current studies was to examine ILM using the cancelation procedure to determine if an individual differences approach would provide insight into whether or not the illusory motion produced in various experimental protocols warrant a common explanation. The logic of the reasoning is that if two conditions produce an illusion of motion for the same underlying reason then individuals who show a large illusion in one condition also should show a large illusion in the other, whereas those who show a small illusion in one condition would likewise show a small illusion in the other. In contrast, if two display protocols evoke illusions of motion for different reasons then there is no expectation beyond chance coincidence why having a large illusion in one condition should systematically pair with a large illusion in the other.

Of course, for any set of findings there can be an infinite number of possible explanations (Popper, 1968) but simply because alternative hypothesis for or against such a correlation may be generated does not invalidate the need to determine if these relationships exist. Based on the pattern of results found

in the studies presented, all combinations of brightening and dimming boxes and appearing and disappearing bars all produce an illusion of motion away from the luminance change provided the change in the bar occurs after the luminance change. Based on the logic of the current analysis approach, therefore, one explanation would be expected to account for the fact that the illusion is away from the luminance change for both onset and offset bars.

It was found that if the bar onsets during the flash, no ILM occurs; however, rILM is found if the bar offsets during the flash. The rILM was not correlated with ILM, meaning knowing that a participant produced large ILM in the offset bar after the flash condition told us nothing about what to expect for the magnitude of their rILM in the offset bar during the flash condition. Due to this apparent lack of a relationship, the suggestion is that rILM arises for reasons that are different from those that result in ILM.

Possible explanations for ILM

The fact that ILM was generated away from a box that became darker suggests that the ILM generated in these displays is not simply due to the spreading of subthreshold activation (Jancke et al., 2004) that is greater around the brighter of the two boxes, because this would predict motion towards the dark flash and away from the brighter nonflashed box. Moreover, there would be an additional reversal for offset bars as well as the increase in subthreshold activity would act to preserve the existing bar rather than hasten the detection of the offset. ILM for offset bars also argues against temporal extensions being generated by exogenous visual attention (Schmidt & Klein, 1997). The findings are, however, consistent with the attentional gradient explanation as this predicts that both bright and dark flashes should attract exogenous attention, which in turn should quicken the detection of both onset and offset events.

There was no evidence for dark flashes to generate stronger illusions in offset bars than onset bars with the opposite pattern for bright cues. The lack of this predicted interaction argues against an impletion mechanism (Downing & Treisman, 1997) that interprets a reduction in luminance and the disappearance of the bar as being more real-world plausible than an increase in luminance and the disappearance of the bar.

Both dark and bright flashes generated stronger illusions in onset bars than offset bars as measured by both ILM_{area} and ILM_{PSE}. While not definitive, the weaker illusion in the offset bars could be viewed as consistent with the suggestion that there are at least two processes that result in ILM (von Grünau, Dube, et al., 1996): a process that generally results in a strong illusion away from either a bright or a dim flash and that is reflected by the strong test-retest reliability. However, when the box darkens, it is the nonflashed box that would be expected to be generating stronger subthreshold activity (Jancke

et al., 2004), and this activity should work to sustain an existing stimuli resulting in the offset line appearing to shoot into the dim flash box and therefore weakens the illusory percept.

Alternatively, the weaker motion signal that results for offset bars could reflect the fact that exogenous attention is thought to spread over a cued object (Egly, Driver, & Rafal, 1994). The argument goes that during offset bar trials when one box is cued attention would spread along the bar reducing the gradient, and therefore reducing ILM. Notwithstanding the importance of these issues, they are beyond the scope of the current investigation and would require careful equating of the real motion signals generated by onset and offset bars (see Crawford et al., 2010 footnote 1).

Possible explanations for rILM

The findings from both onset and offset bars after the flash conditions are easily explainable by a gradient of prior entry benefits that occur after the orienting of exogenous attention to a peripheral luminance change. The failure to obtain ILM for onset bars during the flash also could fit within this framework by arguing that there is insufficient time to orient to the flash and establish the attentional gradient and so the necessary condition of attention being focused at the location has not been met. That, however, makes rILM in offset bars during the flash difficult to explain by the attentional gradient, but the individual difference results indicate that it does not have to be. Rather, rILM in offset bars during the flash appears to be unrelated to ILM after the flash, and this suggests that it arises for a different reason.

If rILM is not attention-based, what explanations might account for it? At the point when the bar is fully removed, the display is comprised of one box brighter than the previously existing bar and another box that was the same luminance. With rILM, the bar appears to have disappeared into the brighter box. If the increased brightness results in a spreading of subthreshold activity (Jancke et al., 2004), then this increased activity would act to sustain the neural activity associated with the previously existing bar. As a result, an illusion towards the brighter box could be explained by the spreading of subthreshold activity.

However, such activity also should have served to quicken the detection of the onset bars in Experiment 2 and so produce forward ILM, which the results do not support. It is reasonable to speculate that the onset of the bar was sufficiently above threshold that the benefit of this activity was not detected. Therefore, if spreading subthreshold activation is the explanation for offset rILM then the prediction is that one would expect to find ILM for near threshold onset bars. Moreover, this ILM should be negatively correlated with the rILM of offset bars with neither correlated with the ILM found after the flash with super-threshold bars. One point to note about

this prediction is that there could be a positive correlation with ILM for near threshold onset bars for a short period of time after the flash, because this situation may show ILM influenced by both attention and the subthreshold activity, but the latter influence should fade quickly, leaving only the attentional influence.

Can the results be explained in terms of contrast counter-change? First, let us start with the ILM findings. When there is no flash and the bar onsets, the bar creates an increase in contrast with the background and there is a decrease in contrast along the edge of the bar where it touches the boxes. This results in motion towards the middle, with no directional bias (Hock & Nichols, 2010). When an existing bar offsets between two boxes, there is a large surface contrast decrease between the bar and the background while there are contrast increases along the edge of the boxes that are revealed. This results in a motion percept from the bar centre outwards towards the two boxes, again without directional bias (Hock & Nichols, 2010).

However, when one box flashes, either getting brighter or darker, the initial change would result in a large surface contrast increase at the flash location. The return to the starting luminance would, therefore, provide a large contrast decrease. Upon returning to the starting luminance, the display becomes that described for the no-flash condition above, in which there is no directional bias in motion perception. Because the results of Experiment 1 indicate motion away from the flash for both onset and offset bars, it must be the decrease of the second luminance change that drives the motion percept as being away from the flash towards either the surface contrast increase created by an onset bar or by the edge contrast increase created by the offset bar.

If the bar changes during the flash, then the flashed box becomes a large contrast increase rather than a decrease. Given that there is a bar related contrast decrease for onset bars at the edges and offset bars over the surface, then the motion should be towards the flash-related increase. In other words, the same situations occur as with ILM except the contrast increase and decrease occur at different locations with respect to the flash location predicting rILM. While the predicted rILM was not found for onset bars, it was found for offset bars. However, if contrast counter-change is driving both ILM and rILM, then these conditions should be related to each other because the same situation, simultaneous increase and decrease in contrast, is proposed to drive both motions. Whereas strong correlations were found for various forms of ILM, and rILM also showed good test-retest reliability in Experiment 3 between the 2AFC and 3AFC versions, there was not the predicted negative correlation between ILM and rILM. Moreover, the data from ILM_{area} and ILM_{con} were both positively in favour of the null hypothesis, although ILM_{PSE} was weakly in favour of a relationship. On the whole, therefore, the data do not favour contrast counter-change

being a common explanation for ILM and rILM. This does not imply, nor should it be taken to mean, that contrast counter-change does not result in illusions of motions, only that the data do not currently suggest that it is the sole explanation for both ILM and rILM. Moreover, it must be remembered that ILM has been reported away from the location of exogenous attention when captured by nonvisual cues (Shimojo, Miyauchi, & Hikosaka, 1997) producing situations that cannot be accounted for by contrast counter-change. As attention cannot account for rILM in the offset bars, it may be that this is a result of contrast counter-change, although the lack of illusory motion for onset bars is troubling for such a conclusion.

It must be remembered that because the rILM that occurs for offset bars has only been shown in two studies, namely the current Experiment 3 and in the Hamm et al. (2014) study, little is known about what aspects of the display change are responsible for its occurrence. We do not know the following: 1) is rILM tied to a process that is time-locked to the onset of the luminance increase or is it a process that is tied to the bar offset—flash offset temporal interval, meaning the 33.7-ms interval following complete bar removal and the offset of the flash; 2) does rILM occur if the flash is a luminance decrease; 3) does rILM occur if the flash does not offset at all; 4) does rILM occur in onset bars which are closer to threshold for detection; 5) does rILM occur if the existing bar changes colour rather than disappears; 6) does rILM occur if the existing bar is not the same luminance and/or colour as the initial boxes; 7) does rILM occur if the bar is not initially touching the boxes; 8) can/does the process that underlies rILM operate simultaneously with processes that produce ILM after the flash?

Before one can suggest with any certainty what mechanism or process underlies rILM in offset bars a more complete understanding of the conditions that produce and influence rILM is required. As noted by Sir Conan Doyle's character Sherlock Holmes in *A Scandal in Bohemia*, "It is a capital mistake to theorize before one has data. Insensibly one begins to twist facts to suit theories, instead of theories to suit facts." (Conan 1892). The current results, however, do suggest that the explanation for rILM in offset bars during the flash differs from that of ILM in onset and offset bars after the flash.

In summary, the current study suggests that examining individual differences in the measures of ILM derived from the cancelation procedure is a promising approach to understanding the mechanisms and processes underlying illusions of motion. In addition, it is suggested that ILM post flash and rILM during the flash, at least with offset bars, arise for different underlying reasons. Whereas the findings for ILM are consistent with the attentional gradient hypothesis, the findings for rILM are not. However, rILM may reflect motion that results from contrast counter-change or possibly from the spreading of subthreshold activity. Both of these explanations, however,

struggle with the fact that no ILM of any sort was found with onset bars presented during the flash, and it may be that rILM requires a different explanation altogether.

Appendix A

It was noted in Crawford et al. (2010) that when ILM and real motion were in the same direction that decision times were faster than when ILM and real motion conflicted. We examined the possibility of using this difference as another index of the illusion, speculating that those who showed a larger illusion might therefore show a larger difference in their response times between conflicting and congruent motions. This quantity will be referred to as the decision time congruency effect and is calculated as the mean of the congruent trials subtracted from the mean of the incongruent trials. Congruency of motion signals is based on the assumption that ILM is away from the flashed location, resulting in a negative value if the motion were towards the flash. Specifically, then the congruent trials are all left real motion trials paired with a right flash and all right real motion trials paired with the left flash. In congruent trials are all left real motion trials paired with a left flash and all right real motion trials paired with a right flash.

Decision time congruency effect

The congruency effect was analysed in a two-way repeated measures ANOVA, with bar type (appear vs. disappear) and cue (bright vs. dim) as factors. Mean congruency effects were 180, 203, 170, and 168 ms for bar appears with bright cues, bar appears with dim cues, bar disappears with bright cues, and bar disappears with dim cues, respectively. Standard deviations were 121, 147, 89, and 98, respectively. The analysis indicated there was no main effect of bar type ($F(1,22) = 0.553$, $p = 0.465$, $\eta^2 = 0.025$). There was no main effect of cue ($F(1,22) = 1.098$, $p = 0.306$, $\eta^2 = 0.048$). There was no bar type by cue interaction ($F(1,22) = 1.074$, $p = 0.311$, $\eta^2 = 0.047$).

After collapsing all the conditions together, the mean congruency effect was calculated for each participant. After removing any data pairs with excessive Cook's distance values, the congruency effect was correlated with ILM_{area}, ILM_{PSE}, and ILM_{con} (results in Fig. 19, A-C). None of the relationships reached significance at the Bonferroni corrected level of significance of $p < 0.0167$ with $r_{(21)} = 0.1824$, $p = 0.4048$, $p_{H0|D} = 0.7646$, $r_{(22)} = 0.1091$, $p = 0.6118$, $p_{H0|D} = 0.8093$, and $r_{(22)} = 0.3937$, $p = 0.0570$, $p_{H0|D} = 0.3937$, for the relationship between the congruency effect and ILM_{area}, ILM_{PSE}, and

ILM$_{con}$, respectively. These are considered positive evidence in favour of the null hypothesis with respect to ILM$_{area}$ and ILM$_{PSE}$, although there is weak evidence against the null hy-

pothesis with respect to ILM$_{con}$ (Raftery, 1995). If the one excessive data point is included for the prediction of ILM$_{area}$, the r value is 0.07).
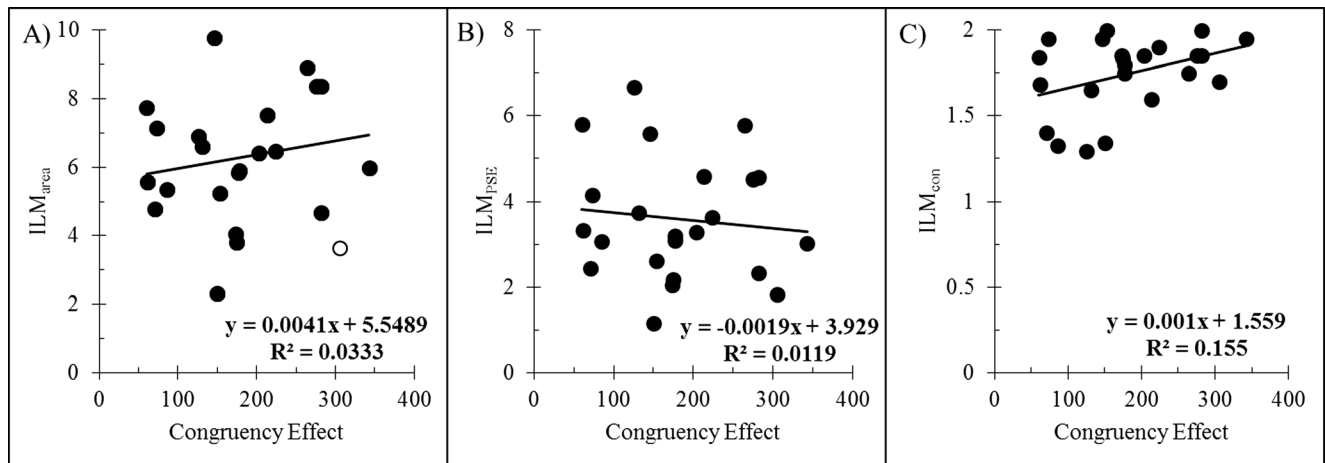

Fig. 19 Correlations between the decision time congruency effect and (A) ILM$_{area}$, (B) ILM$_{PSE}$, and (C) ILM$_{con}$ after collapsing over bright and dim cues for both onset and offset bars. Open circles indicate data pairs omitted due to excessive Cook's distance values

In summary, at the group level analysis the congruency effect does not show a similar pattern as was found in the percept scores. Moreover, the congruency effect does not correlate with the percept measures of ILM, although there is a suggestion of a relationship with the ILM$_{con}$ measure. It seems, therefore, that the congruency effect may not reflect the strength of the illusion per se or it may be that the congruency effect is simply more unstable and requires a much larger number of trials to stabilize. This seems unlikely, because the congruency effect in the final correlations were based on the difference between two mean response times of 240 trials each.

Appendix B

Decision time congruency effect

The congruency effect for the decision times was calculated. It should be noted that in all conditions, it is calculated on the assumption that ILM would be away from the flash, which means negative values are returned for rILM. The congruency effect was analysed in a two-way repeated measures ANOVA, with bar condition (after vs. during the flash) and response options (2AFC vs. 3AFC) as factors. Mean congruency effects were 93, 123, −24, and −25 ms for bar after the flash with 2AFC, bar after the flash with 3AFC, bar during the flash with 2AFC, and bar during the flash with 3AFC, respectively. Standard deviations were 47, 47, 22, and 28, respectively.

The analysis indicated there was a main effect of bar condition ($F_{(1,22)} = 230.541$, $p < 0.001$, $\eta^2 = 0.909$). There was a main effect of response options ($F_{(1,22)} = 8.080$, $p = 0.009$, $\eta^2 = 0.260$). There was an bar condition by response option interaction ($F_{(1,22)} = 5.450$, $p = 0.029$, $\eta^2 = 0.192$).

As shown in Fig. 20 panels A-C, in the bar after the flash conditions, the congruency effect did not correlate with ILM$_{area}$, ILM$_{PSE}$, or ILM$_{con}$ ($r_{(20)} = 0.04$, $p = 0.860$, $p_{H0|D} = 0.8217$, $r_{(20)} = −0.11$ $p = 0.626$, $p_{H0|D} = 0.8040$, and $r_{(21)} = 0.40$, $p = 0.059$, $p_{H0|D} = 0.3924$, respectively). These findings for ILM$_{area}$ and ILM$_{PSE}$ are both considered positive evidence in favour of the null hypothesis, whereas ILM$_{con}$ could be considered as weak evidence against the null hypothesis. However, the congruency effect was correlated to ILM$_{area}$, ILM$_{PSE}$, and ILM$_{con}$ in the bar during the flash conditions (panels D-F: $r_{(20)} = 0.79$, 0.79, and 0.70, all $p < 0.001$, $p_{H0|D} < 0.003$), all which produced very strong evidence against the null. Fischer's z-transformation (Cohen & Cohen, 1983) indicated that the relationship was significantly (Bonferroni corrected criterion of $p < 0.0167$) stronger in the bar during the flash condition for ILM$_{area}$ ($z = 3.25$, $p < 0.001$), ILM$_{PSE}$ ($z = 3.04$, $p = 0.001$), but not ILM$_{con}$ ($z = 1.37$, $p = 0.088$). The relationships if the outlier data pairs are included are $r_{(22)} = 0.07$, −0.06, and 0.28 for the ILM$_{area}$, ILM$_{PSE}$, and ILM$_{con}$ in the bar after the flash conditions, and $r_{(22)} = 0.60$, 0.54, and 0.75, respectively, in the bar during the flash condition, with the same pattern of significance based on the Fischer's z-transformation comparisons.
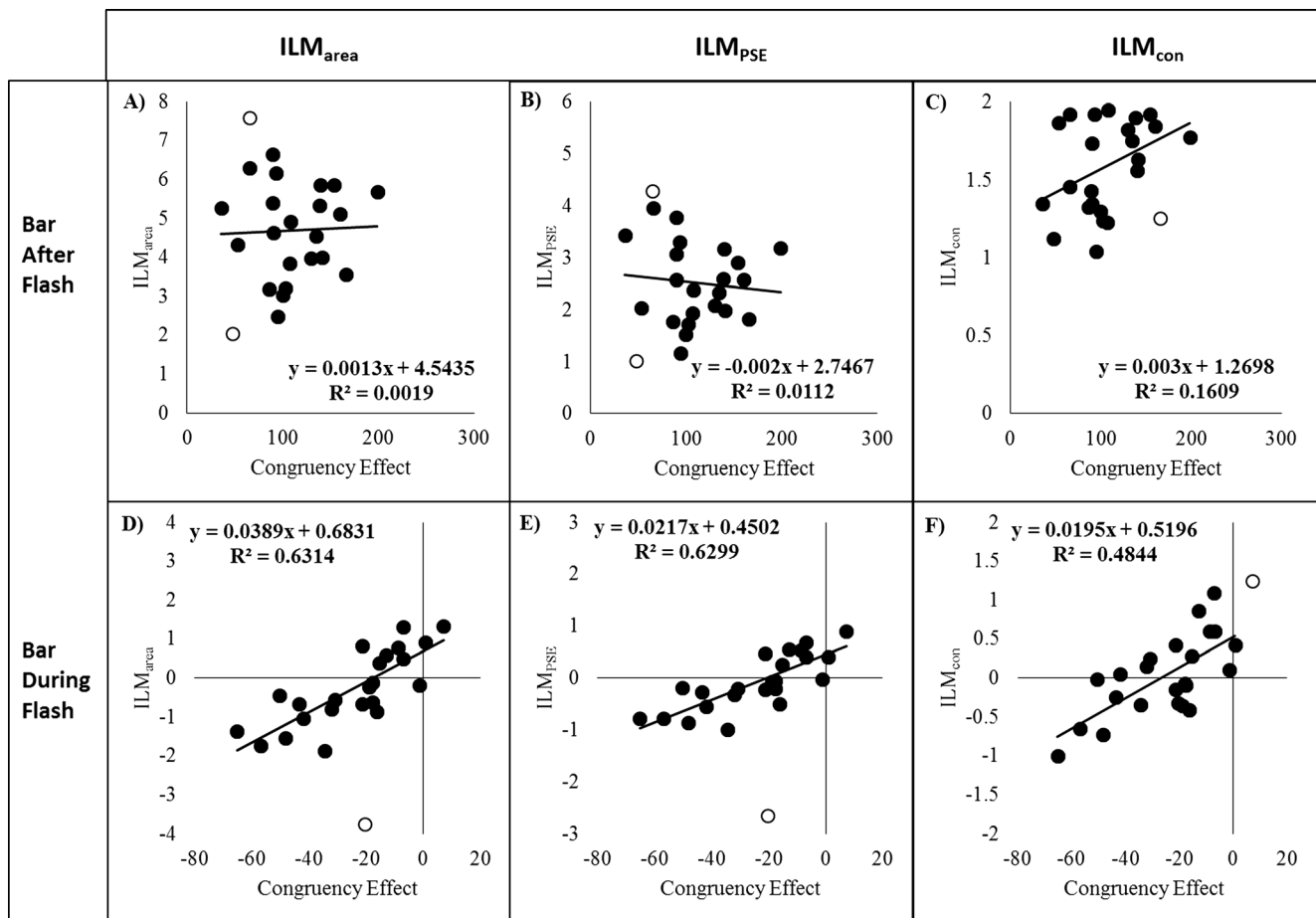
Fig. 20 Correlations between the decision time congruency effect and (A) ILM_{area}, (B) ILM_{PSE}, and (C) ILM_{con} for the bar after the flash condition and (D) ILM_{area}, (E) ILM_{PSE}, and (F) ILM_{con} for the bar during the flash condition. Open circles indicate data pairs omitted due to excessive Cook's distance values

It is of interest that the congruency effect was correlated with the percept measures in the absence of any illusion but neither ILM_{area} nor ILM_{PSE} were correlated with the congruency effect when ILM did occur. ILM_{con} tended to be correlated with the congruency effect in the presence of ILM and was correlated with ILM_{con} in the absence of an illusion of motion. Similar to Experiment 1, the suggestion is that the congruency effect may not index the illusion's strength but may reflect other aspects of the task or decision making process.

Appendix C

Decision time congruency effect

The congruency effect was analysed in a two-way repeated measures ANOVA, with bar condition (after vs. during the flash) and response options (2AFC vs. 3AFC) as factors. Mean congruency effects were 88, 111, −46, and −53 ms for

bar after the flash with 2AFC, bar after the flash with 3AFC, bar during the flash with 2AFC, and bar during the flash with 3AFC, respectively. Standard deviations were 67, 59, 30, and 28, respectively. The analysis indicated there was a main effect of bar condition ($F_{(1,22)} = 97.958$, $p < 0.001$, $\eta^2 = 0.810$). There was a main effect of response options ($F_{(1,22)} = 4.621$, $p = 0.042$, $\eta^2 = 0.167$). There was an bar condition by response option interaction ($F_{(1,22)} = 11.503$, $p = 0.003$, $\eta^2 = 0.333$).

The decision time congruency effect was analysed to determine whether it was correlated with ILM_{area}, ILM_{PSE}, and ILM_{con} for both the bar after the flash and bar during the flash conditions to assess if the congruency effect was related to either ILM or rILM. The Bonferroni corrected criterion for significance was $p < 0.0167$. The scatter plots are shown in Fig. 21 C1, A-F. As with other cases of ILM shown in Experiments 1 and 2, in the bar after the flash condition the congruency effect was not significantly correlated with ILM_{area} ($r_{(21)} = 0.18$, $p = 0.411$, $p_{H0|D} = 0.7666$, positive evidence in favour of the null hypothesis), ILM_{PSE} ($r_{(20)} = −0.01$,

$p = 0.965$, $p_{H0|D} = 0.8273$, positive evidence in favour of the null hypothesis), or ILM$_{con}$ ($r_{(20)} = 0.45$, $p = 0.036$, $p_{H0|D} = 0.2802$, weak evidence against the null hypothesis). However, in the bar during the flash condition, the relationship just failed to exceed the criterion for significance for ILM$_{area}$ ($r_{(21)} = 0.49$, $p = 0.017$, $p_{H0|D} = 0.1694$) but was significant for ILM$_{PSE}$ ($r_{(20)} = 0.54$, $p = 0.010$, $p_{H0|D} = 0.0956$) and ILM$_{con}$ ($r_{(21)} = 0.51$, $p = 0.013$, $p_{H0|D} = 0.1305$). All three of these latter findings would be considered positive evidence against

the null hypothesis. While the pattern suggests an interaction, none of the comparisons between the bar conditions reached significance (Bonferroni corrected criterion of $p < 0.0167$), ILM$_{area}$ ($z = 1.13$, $p = 0.258$), ILM$_{PSE}$ ($z = 1.81$, $p = 0.070$), and ILM$_{con}$ ($z = 0.25$, $p = 0.803$). Including the outlier data pairs results in $r = 0.30$, 0.13, and 0.33 for the bar after the flash measures of ILM$_{area}$, ILM$_{PSE}$, and ILM$_{con}$, respectively, and $r = 0.61$, 0.58, and 0.45 for the bar during the flash measures.
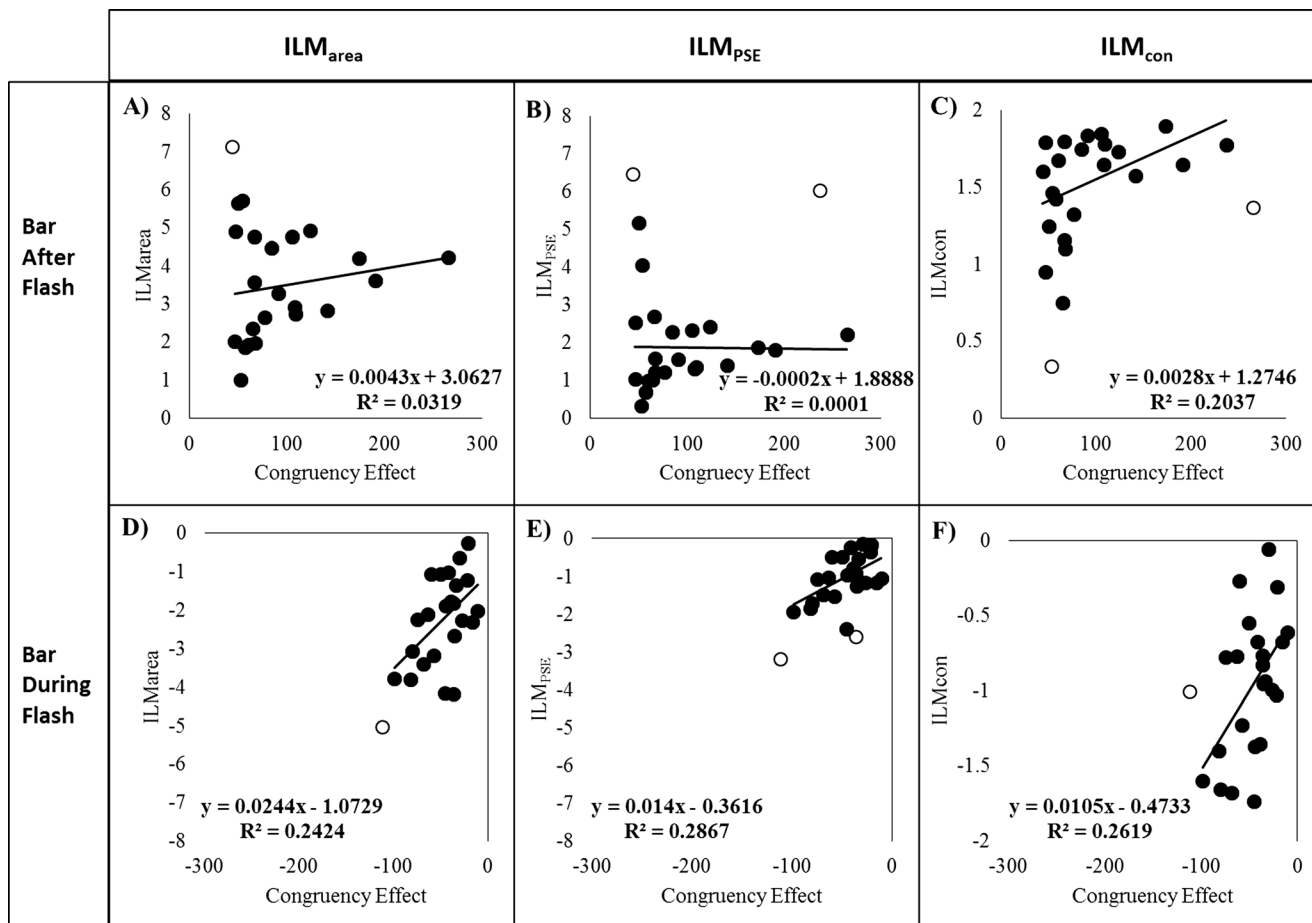

Fig. 21 Upper row depicts the correlations during the bar after the flash condition between the decision time congruency effect and (A) ILM$_{area}$, (B) ILM$_{PSE}$, and (C) ILM$_{con}$. Lower row depicts the correlations during the bar during the flash condition between the decision time congruency effect and (D) ILM$_{area}$, (E) ILM$_{PSE}$, and (F) ILM$_{con}$. Open circles indicate data pairs omitted due to excessive Cook's distance values

Similar to previous findings, the congruency effect was not related to ILM$_{area}$ or ILM$_{PSE}$ and the relationship with ILM$_{con}$ was tentative, failing to survive the Bonferroni cutoff for significance. In contrast, the congruency effect was tentatively correlated with rILM$_{area}$, whereas the correlation with ILM$_{PSE}$ and ILM$_{con}$ both survived the Bonferroni correction. Again, it appears that the congruency effect is not an index of the illusion per se but may reflect some other process being

influenced in this task. The fact that rILM and the congruency effect appear to be related is similar to the finding in Experiment 2 where the congruency effect was related to the percept measures obtained with then line changed during the flash. This again suggests that ILM and rILM may not be two versions of the same illusion, although it should be noted the tests between the correlations do not support the assertion of an interaction. Decision times will be influenced by many

factors in addition to the perception of motion (Jensen, 2006). That means that the relationships between the congruency effect and ILM/rILM measures may have to do with aspects of the display or task other than the mechanisms responsible for the perception of motion. This suggests that these tasks should be viewed as being rich in psychological content worthy of systematic study.

References

Bavelier, D., Schneider, K. A., & Monacelli, A. (2002). Reflexive gaze orienting induces the line-motion illusion. *Vision Research, 42,* 2817–2827.

Cartwright, D. (1941). Relation of decision time to the categories of response. *American Journal of Psychology, 54,* 174–196.

Chica, A. B., Lasaponara, S., Lupianez, J., Doricchi, F., & Bartolomeo, P. (2010). Exogenous attention can capture perceptual consciousness: ERP and behavioural evidence. *NeuroImage, 51*(3), 1205–1212. doi:10.1016/j.neuroimage.2010.03.002

Christie, J. (2014). Illusory line motion is not caused by object-differentiating mechanisms or endogenous attention. *Quarterly Journal of Experimental Psychology, 67*(12), 2293–2300. doi:10.1080/17470218.2014.918633

Christie, J., & Barresi, J. (2002). Using illusory line motion to differentiate misrepresentation (Stalinesque) and misremembering (Orwellian) accounts of consciousness. *Consciousness and Cognition, 11*(2), 347–365. doi:10.1006/ccog.2001.0539

Christie, J., & Klein, R. M. (2005). Does attention cause illusory line motion? *Perception & Psychophysics, 67*(6), 1032–1043.

Cohen, J., & Cohen, P. (1983). *Applied multiple regression/correlation analysis for the behavioral sciences* (2nd ed.). Hillsdale: Erlbaum.

Conan Doyle, A. (1892). *The adventures of Sherlock Holmes.* United Kingdom: George Newnes

Corballis, P. M., Funnell, M. G., & Gazzaniga, M. S. (2002). An investigation of the line motion effect in a callosotomy patient. *Brain and Cognition, 48*(2-3), 327–332. doi:10.1006/brcg.2001.1371

Crawford, T. J., Hamm, J. P., Kean, M., Schmechtig, A., Kumari, V., Anilkumar, A. P., & Ettinger, U. (2010). The perception of real and illusory motion in schizophrenia. *Neuropsychologia, 48*(10), 3121–3127. doi:10.1016/j.neuropsychologia.2010.06.027

Crawford, T. J., Kean, M., Klein, R. M., & Hamm, J. P. (2006). The effects of illusory line motion on incongruent saccades: Implications for saccadic eye movements and visual attention. *Experimental Brain Research, 173,* 498–506. doi:10.1007/s00221-006-0392-z

Downing, P. E., & Treisman, A. M. (1997). The line-motion illusion: Attention or impletion? *Journal of Experimental Psychology: Human Perception & Performance, 23*(3), 768–779.

Egly, R., Driver, J., & Rafal, R. D. (1994). Shifting visual attention between objects and locations: Evidence from normal and parietal lesion subjects. *Journal of Experimental Psychology. General, 123*(2), 161–177.

Faubert, J., & von Grünau, M. (1995). The influence of two spatially distinct primers and attribute priming on motion induction. *Vision Research, 35*(22), 3119–3130. doi:10.1016/0042-6989(95)00061-4

Hamm, J. P. (2001). Object-oriented millisecond timers for the PC. Behavior Research Methods Instruments & Computers, 33(4), 532–539. doi:10.3758/BF03195412

Hamm, J. P., Crawford, T. J., Nebl, H., Kean, M., Williams, S. C., & Ettinger, U. (2014). Neural correlates of illusory line motion. *PloS One, 9*(1), e87595. doi:10.1371/journal.pone.0087595

Hamm, J. P., & Klein, R. M. (2002). Does attention follow the motion in the "shooting line" illusion? *Perception & Psychophysics, 64*(2), 279–291. doi:10.3758/BF03195792

Heathcote, A. (1988). Screen control and timing routines for the IBM microcomputer family using a high-level language. *Behavior Research Methods, Instruments and Computers, 20,* 289–297

Hikosaka, O., Miyauchi, S., & Shimojo, S. (1993a). Focal visual attention produces illusory temporal order and motion sensation. *Vision Research, 33*(9), 1219–1240.

Hikosaka, O., Miyauchi, S., & Shimojo, S. (1993b). Visual attention revealed by an illusion of motion. *Neuroscience Research, 18*(1), 11–18.

Hikosaka, O., Miyauchi, S., & Shimojo, S. (1993c). Voluntary and stimulus-induced attention detected as motion sensation. *Perception, 22,* 517–526.

Hikosaka, O., Miyauchi, S., Takeuchi, H., & Shimojo, S. (1996). Multimodal spatial attention visualized by motion illusion. In T. Inui & J. L. McClelland (Eds.), *International Symposium on Attention and Performance* (pp. 237–261). London: MIT Press.

Hock, H. S., & Nichols, D. F. (2010). The Line Motion Illusion: The Detection of Counterchanging Edge and Surface Contrast. *Journal of Experimental Psychology-Human Perception and Performance, 36*(4), 781–796. doi:10.1037/A0016876

Jancke, D., Chavane, F., Naarman, S., & Girinvald, A. (2004). Imaging cortical correlates of illusion in early visual cortex. *Nature, 428,* 423–426.

Jensen, A. R. (2006). *Clocking the mind: Mental chronometry and individual differences* (1st ed.). Amsterdam: Elsevier.

Jonides, J., & Mack, R. (1984). On the cost and benefit of cost and benefit. *Psychological Bulletin, 96*(1), 29–44. doi:10.1037/0033-2909.96.1.29

Kanizsa, G. (1979). *Organization in vision: Essays on Gestalt perception.* New York: Praeger.

Kraepelin, E. (1919). *Dementia Praecox and Paraphrenia (R.M. Barclay, Trans.).* Edinburgh: E. & S. Livingstone.

Laberge, D. (1983). The spatial extent of attention to letters and words. *Journal of Experimental Psychology: Human Perception & Performance, 9,* 371–379.

Masson, M. E. J. (2011). A tutorial on a practical Bayesian alternative to null-hypothesis significance testing. *Behavior Research Methods, 43*(3), 679–690. doi:10.3758/s13428-010-0049-5

Masson, M. E. J., & Loftus, G. R. (2003). Using confidence intervals for graphically based data interpretation. *Canadian Journal of Experimental Psychology, 57*(3), 203–220.

Myors, B. (1999). Timing accuracy of PC programs running under DOS and Windows. *Behavior Research Methods, Instruments, & Computers, 31*(2), 322–328.

Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia, 9,* 97–113.

Popper, K. R. (1968). *The logic of scientific discovery* (revisedth ed.). London: Hutchinson & Co (LTD).

Posner, M. I., & Cohen, Y. (1984). Components of visual orienting. In H. Bouma & D. Bowhuis (Eds.), *Attention and Performance X* (pp. 531–556). London: Erlbaum.

Raftery, A. E. (1995). Bayesian model selection in social research. *Sociological Methodology, 25*(25), 111–163. doi:10.2307/271063

Schmidt, W. C., & Klein, R. M. (1997). A spatial gradient of acceleration and temporal extension underlies three illusions of motion. *Perception, 26*(7), 857–874.

Shimojo, S., Miyauchi, S., & Hikosaka, O. (1997). Visual motion sensation yielded by non-visually driven attention. *Vision Research, 37*(12), 1575–1580. doi:10.1016/S0042-6989(96)00313-6

Shore, D. I., Spence, C., & Klein, R. M. (2001). Visual prior entry. *Psychological Science, 12*(3), 205–212.

Steinman, B. A., Steinman, S. B., & Lehmkuhle, S. (1995). Visual attention mechanisms show a centre-surround organization. *Vision Research, 35*(13), 1859–1869.

Tanabe, H. C., & Yanagida, T. (2002). *Neural activity correlated with intra- or cross-modal cued illusory line motion: an fMRI study*. Paper presented at the Human PSI Forum, OVTA, Makuhari, Chiba, Japan

von Grünau, M., Dube, S., & Kwas, M. (1996). Two contributions to motion induction: A preattentive effect and facilitation due to attentional capture. *Vision Research, 36*(16), 2447–2457.

von Grünau, M., & Faubert, J. (1994). Intraattribute and interattribute motion induction. *Perception, 23*(8), 913–928. doi:10.1068/p230913

von Grünau, M., Racette, L., & Kwas, M. (1996). Measuring the attentional speed-up in the motion induction effect. *Vision Research, 36*(16), 2433–2446.

Zanker, J. M. (1997). Is facilitation responsible for the "motion induction" effect? *Vision Research, 37*(14), 1953–1959. doi:10.1016/S0042-6989(97)00004-7