# Joint attention enhances visual working memory

Samantha EA Gregory & Margaret C Jackson

School of Psychology, University of Aberdeen

AUTHORS' COPY

Running head: Joint attention working memory

Key Words

Joint attention
Gaze Cuing
Working memory
Attention Orienting

Correspondence to:
Samantha Gregory
School of Psychology
University of Aberdeen
Aberdeen
AB24 3FX
s.gregory@abdn.ac.uk

## Abstract

Joint attention - the mutual focus of two individuals on an item – speeds detection and discrimination of target information. But what happens to that information beyond the initial perceptual episode? To fully comprehend and engage with our immediate environment also requires working memory (WM), which integrates information from second to second to create a coherent and fluid picture of our world. Yet no research exists at present that examines how joint attention directly impacts WM. To investigate this, we created a unique paradigm which combines gaze cues with a traditional visual WM task. A central, direct gaze ‘cue’ face looked left or right, followed 500ms later by 4, 6, or 8 coloured squares presented on one side of the face for encoding. Crucially, the cue face either looked at the squares (valid cue) or looked away from them (invalid cue). A no shift (direct gaze) condition served as a baseline. After a blank 1000ms maintenance interval, participants stated whether a single test square colour was present or not in the preceding display. WM accuracy was significantly greater for colours encoded in the valid versus invalid and direct conditions. Further experiments showed that an arrow cue and a low-level motion cue - both shown to reliably orient attention - did not reliably modulate WM, indicating that social cues are more powerful. This study provides the first direct evidence that sharing the focus of another individual establishes a point of reference from which information is advantageously encoded into WM.

Humans display a strong sensitivity to others' eye gaze. From birth infants prefer to look at faces that are looking at them (Farroni, Csibra, Simion, & Johnson, 2002), from as young as 3 months they begin to follow the gaze of others (e.g., Hood, Willen, & Driver, 1998), by 8 months they can selectively use gaze cues so that only reliable information is utilised (Tummeltshammer, Wu, Sobel, & Kirkham, 2014), and by 3 years can make reliable judgements of exactly what a person is looking at (Doherty, 2006). Gaze following develops gradually into what has been termed 'joint attention', the shared focus of two individuals on an object, person, or event. The ability to engage in joint attention is linked to theory of mind, the knowledge that others hold a different perspective and different knowledge of the world than we ourselves do (e.g. Charman et al., 2000). This develops further into mentalization skills, which specifically relate to the way in which humans make inferences about the mental states of others, allowing an individual to take another's perspective (Frith & Frith, 2006). Social cues, such as where someone looks, are utilised in making such inferences.

Based on the traditional Posner cuing task (Posner, 1980) designed to measure orienting of spatial attention, a simple yet elegant gaze-cuing paradigm has been developed to investigate joint attention in the laboratory (Friesen & Kingstone, 1998). A central cue face looks left or right and a subsequent target is presented in either the looked at (validly cued) or looked away from (invalidly cued) location. Target tasks typically involve simple location or detection of an asterisk or other item (e.g. Friesen & Kingstone, 1998) or discrimination of a letter (Driver et al., 1999). Task performance is reliably found to be faster and more accurate in the valid versus invalid gaze condition, indicating that the observer has oriented attention to share the focus of the cue face and engaged in joint attention. This gaze cuing effect has a number of characteristics, which imply that it occurs reflexively in an exogenous manner, without top-down volitional control. It can occur at very short cue-target stimulus onset asynchronies (SOA's) with cuing found at 0ms SOA (Xu, Zhang, & Geng, 2011), and observers continue to follow gaze even when the gaze cue is entirely non-predictive of target location and thus gaze following may be detrimental to task performance (Friesen & Kingstone, 1998). Gaze cuing also engages volitional, endogenous attention orienting, thought to occur at SOAs upwards of 300ms. Valid gaze cues are shown to reliably facilitate task performance at SOAs of up to 700ms (Frischen, Bayliss, & Tipper, 2007) despite the fact that observers explicitly understand that the cue is uninformative (50% valid, 50% invalid). However, eye gaze is not unique in guiding spatial attention. Non-social, central, symbolic cues such as arrows and directional words also reliably orient attention across a similar time course, showing both reflexive and voluntary orienting from approximately 100ms SOA to upwards of 600ms (Hommel, Pratt, Colzato, & Godijn, 2001; Ristic, Friesen, & Kingstone, 2002; Taylor & Klein, 2000; Tipples, 2002, 2008). Despite similarities between social and non-social cuing effects, the mechanisms underpinning the orienting of attention to gaze and non-social symbolic cues are thought to be different. Effects elicited by non-social symbolic cues such as arrows are proposed to reflect *automated symbolic attention* due to over-learned connections between the cue and its typical response (Ristic & Kingstone, 2012; Ristic, Landry, & Kingstone, 2012; Ristic & Landry, 2014), whereas the response to gaze cues is thought to reflect a social-biological response (Emery, 2000). Further, using functional imaging, it has been found that that while eye gaze cuing triggers focussed activation related to enhanced visual processing, thus pointing to more reflexive processing, arrow cues activate a much broader network including areas specifically related to volitional orienting (Hietanen, Nummenmaa, Nyman, Parkkola, & Hämäläinen, 2006).

Research into the effect of joint attention on cognition demonstrates that seeing something in the context of another’s gaze can influence perceptual processing beyond the realms of basic attention. For example, Bayliss and colleagues (Bayliss, Frischen, Fenske, & Tipper, 2007; Bayliss, Paul, Cannon, & Tipper, 2006) found that items seen in the context of eye gaze cues were rated as more likeable by participants than objects looked away from by the cues, an effect not found with arrow cues. Further, joint attention has been found to improve language comprehension in children (Tomasello & Farrar, 2015; Tomasello, 1988) and adults (Hanna & Brennan, 2007; Knoeferle & Kreysa, 2012), plays an important role in object learning (e.g., Cleveland, Schug, & Striano, 2007; Striano, Chen, Cleveland, & Bradshaw, 2006; Wu, Gopnik, Richardson, & Kirkham, 2011), and influences spatial learning in LTM for object locations among infants (Wu & Kirkham, 2010). This level of understanding of and engagement with the world around us clearly shows that sharing the attentional focus of another individual extends beyond attention and perception to involve memory. Once information has been attended, located, and defined, it requires involvement of working memory (WM) to contribute to immediate goal-directed behaviour, and long-term memory (LTM) to facilitate learning and retain information for longer-term future reference.

While studies that show an influence of joint attention on learning undoubtedly engage WM and LTM at various stages in the process, there is scant literature that isolates these two forms of memory in the absence of a more complex interactive learning task. Regards LTM, Richardson and colleagues (2012) utilised a ‘looking together’ paradigm where participants were informed that another participant, with whom they did not interact, could either see what they could see or saw something else. They found that when participants were told they were looking at an image with another who was engaged in the same task, recall of the image was greater than when they were told that the other was looking at a different image.  This study shows that imagined joint attention can enhance LTM relative to independent thought. Utilising a similar paradigm, Shteynberg, (2010) specifically looked at the effect of similar vs dissimilar individuals, finding that LTM enhancements occurred only when the imagined joint attention was with a similar other. Kim and Mundy (2012) examined how memory for images was influenced by initiated versus followed gaze. They found better long-term recognition memory for images encoded under joint attention, but note that in this study participants were explicitly told to follow the gaze cues so this does not explore natural tendencies to utilise (or ignore) social cues.

Also investigating the influence of joint attention on LTM, and utilising the traditional gaze cuing paradigm with uninformative cues, Dodd, Weiss, McDonnell, Sarwal, and Kingstone (2012) showed common object words presented to the left or right of a central, schematic cue face. Before each word was presented, the face looked left or right and 500ms later a word appeared in either the looked at or the looked away from location. Participants were informed that gaze direction was not predictive of where the word would appear (50% valid, 50% invalid). After all items had been presented, participants were required to write down all the words they could recall. Dodd and colleagues found that more words were recalled when they were encoded in the context of valid, shared gaze (i.e., encoded under joint attention) than words presented in the invalid condition, both when encoding was intentional (Experiment 2) and incidental (Experiment 3). It is important to note that Dodd and colleagues' suggest that their data points to a detriment for the looked away from items rather than enhanced memory for looked at items. However, they did not include a direct gaze baseline, so it is difficult to unpack this. Interestingly, arrow cues did not exert a significant effect on word recall, implying the effect is social in nature. Taken as a whole, these studies provide evidence that gaze signals can do more than simply orient attention to enhance perception, they can influence LTM in socially relevant ways.

While LTM is useful for embedding information for later use and reference, effective social interaction and information processing relies fundamentally on the engagement of WM processes. Yet there exists no research at present on how joint attention directly influences WM. It is important to examine the isolated effects of joint attention on WM because we need to understand the cognitive stages at which joint attention has an impact during social interaction. WM is an online workspace in which information is temporarily stored and/or manipulated for a few seconds in order to create a coherent and fluent representation of our internal thoughts and external world and is essential for reasoning, comprehension, planning, and learning (Baddeley, 2007). This allows us to achieve the multitude of short-term goals required for successful and efficient cognition and behaviour, such as keeping track of others' intentions and holding in mind important visual, verbal, and spatial information until it is has been utilised or is no longer relevant. WM is limited in capacity, and studies show that we can retain only around four simple item's worth of information in our mind at any one point in time (Cowan, 2001, 2010). Crucially, WM is considered to be connected to attentional control (Kane, Bleckley, Conway, & Engle, 2001), and items that capture attention are shown to enjoy a processing advantage in WM (Awh, Vogel, & Oh, 2006).

The purpose of the current research was to determine whether joint attention enhances WM accuracy. Specifically, we investigated whether WM could be enhanced for items encoded in a jointly attended location compared to those encoded in a location that was not within the mutual focus of the observer and the cue face. To do this, we created a paradigm in which non-predictive gaze cues (50% valid, 50% invalid) were combined with a traditional visual WM task. In Experiment 1, we measured WM for coloured squares as a function of whether they were looked at or looked away from by a central cue face, using a dynamic cue, which shifted gaze from direct to left/right. Baseline trials were also included in which gaze remained direct throughout (non-cuing condition). Based on Dodd et al.'s (2012) findings, we predicted that participants would show better WM for looked at (validly cued) items, compared to both looked away from (invalidly cued) items and those that were not cued (direct gaze). To investigate the social nature of any gaze cuing effects found in Experiment 1, we conducted a further two experiments using non-social central cues. In Experiment 2 we used an arrow cue to determine the role of learned symbolism, and in Experiment 3 we used a low-level motion cue to determine the contribution of reflexive attention to the kinetic aspect of the dynamic gaze shift we used. It is possible that lateral movement of the pupil triggers a motion detector (e.g. Abrams & Christ, 2003) and it is this which guides attention rather than the eyes *per se*.

Individuals with social anxiety tend to avoid the eye region of a face (e.g. Horley, Williams, Gonsalvez, & Gordon, 2004; Moukheiber et al., 2010), making joint attention more difficult. Further, individuals with autism are less likely to look at eyes and engage in joint attention (Charman, 2003; Charman et al., 1997), and healthy individuals who show more autistic-like traits show reduced tendency to follow gaze cues (Bayliss & Tipper, 2005). Thus, in Experiment 1 we used established questionnaires to assess whether levels of social anxiety and autistic-like traits in the healthy adult population may modulate the effect of eye gaze on WM.

## General Methods

### *Participants & Apparatus*

Adult volunteers were recruited from the University of Aberdeen in exchange for course credit or monetary reimbursement. All reported normal or corrected to normal vision, including no colour vision deficiency. The University of Aberdeen ethics committee approved all experiments and all ethical procedures were upheld. Stimuli were presented using E-prime software (Schneider, Eschman, & Zuccolotto, 2002) on a Dell LCD monitor (32-bit true colour; resolution 1280 x 1024 pixels Dell P190S 19 inch 4:3).

### *Stimuli*

Ten different coloured squares (29x 29 pixels, visual angle 0.8° x 0.8°) were used as memoranda (black, blue, brown, green, orange, pink, purple, red, turquoise and yellow). The nature of each cue type (gaze, arrow, motion) is described in detail in the relevant experiment sections below.

*Design & Procedure*

We manipulated cue validity (valid, invalid, direct) and WM load (4, 6, 8 squares) as within-subjects conditions, pseudo-randomised. WM load was varied in order to assess whether any effects of cuing were modulated by task difficulty / resource demands. A further condition manipulated the test probe item at retrieval – present or absent in the preceding display – yielding 18 conditions in total. There were 26 trials in each of the 18 conditions, yielding 468 trials in total for each experiment. On probe present trials (50% of the total), the test square came from the valid location on a third of trials, from the invalid location on another third of trials, and from the left or right hand side equally on the remaining direct (uncued) trials. Thus, cues were entirely non-predictive. The trials were self-initiated and were presented over 6 blocks of 78 trials to allow participants to take breaks if required.

Square colour combinations were selected from each set at random, as was the face identity in the gaze condition. The cue was presented in the centre of the screen and indicated left or right directions equally (or remained direct), and squares appeared equally on the left or right side of the cue, fully counterbalanced. Participants were instructed to look at the centre of the cue at the start of each trial, remember all of the coloured squares, and that the direction of the central cue did not predict where the squares would appear. Ten practice trials preceded the main experimental trials.

Figure 1 illustrates a trial sequence in each experiment (1A – gaze cue; 1B – arrow cue; 1C – motion cue). Each trial began with a fixation-cross displayed for 1000ms, followed by the direct version of the cue in the centre of the screen. After 750ms the cue shifted left/right, or remained direct. After a cue-target SOA of 500ms (Experiment 1 – gaze, and Experiment 2 - arrow) or 150ms (Experiment 3 - motion), coloured squares were presented for 100ms for encoding (as per Luck & Vogel, 1997) on one side of the cue, which remained on screen. Squares were configured in arrays of 4, 6 or 8 to provide WM loads near capacity and exceeding capacity in order to avoid ceiling effects on WM performance (see Figure 2). The squares and cue disappeared for a 1000ms WM maintenance interval consisting only of a central fixation cross. A single probe square was then presented in the centre of the screen and participants indicated via a key press whether this square had been present or not in the immediately preceding array. They had 3000ms in which to respond. We used a limited response window in order to ensure that responses occurred within the bounds of WM Participants could easily perform within this time window, with just 0.3% of trials lost due to time out in Experiment 1, 0.4% in Experiment 2, and 0.3% in Experiment 3. The average response time across experiments was 837ms. Accuracy feedback was provided to participants after every trial, with instruction to keep performance as high as possible. The purpose of this frequent feedback was to keep participants focused on the task, as well as to prevent them from engaging in demand characteristics by consciously choosing to follow(or avoid) cue direction, as this would adversely affect accuracy.

### Pilot study

We conducted a pilot study to verify each cue type using the traditional target location cuing paradigm. In this task, the central (non-directional) version of the cue was presented for 750ms, the cue then signalled left or right for either 150ms, 500ms, or 1000ms (SOA), then a target asterisk appeared in either the valid or invalid location (non-predictive; 50% valid). Participants responded using their left and right index fingers on keyboard keys q (left) and p (right), they were instructed to respond quickly and accurately and that the cue had no bearing on target location. There were 24 trials per condition (144 trials in total). Data from incorrect trials and trials where responses were faster than 200ms (anticipatory) and longer than 1000ms (outliers) were removed (<1% of data).

*Figure 1 about here*

*Figure 2 about here*

## Data Analysis

In all experiments, we conducted statistical analysis on d' values as a measure of WM accuracy, this is an intuitive way of looking at the participants' memory sensitivity as the method combines the results of probe present (Hits) and probe absent trials (False Alarms; FAs), which are both of importance in measuring memory accuracy. Hits and FAs were computed for all three gaze conditions (valid, invalid, direct) and entered into the following formula: d' = z(Hits) – z(FAs). A d' of 0 would indicate no discrimination between the targets, this could be achieved by participants always choosing to say the probe was present, thus they would have a perfect score in the probe present trials, but an accuracy score of 0 in the probe absent trials (i.e., 50% accuracy). Alternatively, a d' of 4.66 would indicate perfect performance in both probe present and probe absent trials (100% accuracy). Therefore, if coloured squares are encoded better in the valid than invalid cue condition, this should lead to more accurate comparison / discrimination between the probe item at retrieval with the contents of WM, yielding higher d' values. Hits, FAs, K (capacity estimates: K = WM Load*(Hits – FAs)) and RTs as a function of WM load and cue validity from each experiment are also provided in Table 1 for reference.

Participant outliers in all three experiments were identified using the median absolute deviation (MAD; Leys, Ley, Klein, Bernard, & Licata, 2013) at the recommended threshold of the median ± 2.5 times the MAD. We used MAD in preference to standard deviation (SD) because outliers can adversely skew calculation of SD but not MAD. Performance at load 4 (the easiest condition) was assessed to identify individuals that performed significantly worse (or better) than the sample. Two participants were excluded from Experiment 1 with d' scores $\leq .11$ (median =1.47; MAD= 0.54); no participants were excluded from Experiments 2 or 3.

For each experiment we conducted a repeated-measures ANOVA on d' values with cue direction (valid, invalid, direct) and load (4, 6, 8) as within-subjects variables. Planned

Bonferroni corrected follow up tests were also conducted where appropriate. Results are reported for each experiment using standard null hypothesis significance testing with additional analysis conducted with Bayesian statistics using JASP (Version 0.7; Love et al., 2015). Bayesian analysis, using equivalent tests (Bayesian repeated measures ANOVAs, and Bayesian paired samples t tests), were conducted on the key main effects related to the gaze cuing effect, as well as planned follow up comparisons. Using JASP's inbuilt interpretation tables and focussing on Bayes factor $BF_{10}$, results can be considered anecdotal evidence that H1 is true when $BF_{10}$ is between 1 and 3, moderate between 3 and 10, and strong evidence above 10. Additionally the likelihood that the null hypothesis is true can be indicated; $BF_{10} = 1$ indicates that the data lends equal support to H1 and H0. Moderate support for H0 is indicated when $BF_{10}$ is between 0.33 and 0.10, and strong evidence for accepting the null is indicated when $BF_{10} \leq 0.10$. This additional analysis is an important tool to check the robustness of findings, as well as helping to understand the nature of the null results.

## Experiment 1: Gaze cue

### Method

#### *Participants*

In order to reliably make any inferences based on individual differences, we recruited a sample of 64 naive adult volunteers (51 female; mean age 20 years).

#### *Stimuli, Design, & Procedure*

We selected a set of faces to use as gaze cues from the Radboud Faces Database (Langner et al., 2010). These comprised six different male and six different female individual identities. Cue face gender is found in other studies to have no effect on participant orienting (see Bayliss, Pellegrino, & Tipper, 2005), so we included both genders in order to increase the ecological validity of the study. Each face identity had three naturally photographed gaze states - left, right, and direct - were cropped to remove hair, and were presented in grey scale.

Faces were presented in the centre of the screen and face dimensions were uniform (109x151 pixels, subtending a visual angle of 3.2°x4.1° at a viewing distance of 60cm). We additionally administered the Liebowitz Social Anxiety Scale (Fresco et al., 2001) and Autism Spectrum Questionnaire (ASQ; Baron-Cohen, Wheelwright, Skinner, Martin, & Clubley, 2001) via Eprime, at the end of the session. Figure 1A illustrates the trial procedure and timing parameters.

The 500ms SOA we used replicates Dodd et al.'s (2012) approach, and falls within the period in which gaze cuing effects are most reliably found (between 300ms and 700ms SOA; Frischen, et al, 2007). A pilot study (see the general methods section) with a sample of 22 naïve and independent participants, confirmed that the specific gaze cues we used here reliably oriented attention to a laterally presented asterisk at the crucial 500ms SOA (valid: M = 337ms; invalid: M = 348ms; $t(21)=2.777$, $p=.011$,Cohen's $d=0.592$, $BF_{10}=4.496$; see figure 3). We also found that gaze cued attention at 150ms (valid: M =364ms; invalid: M = 378ms; $t(21)=2.792$, $p=.011$,Cohen's $d=0.595$; $BF_{10}=4.620$), consistent with other findings (e.g.Friesen & Kingstone, 1998). Interestingly, there was evidence of inhibition of return (IOR) at 1000ms where RTs were faster on invalid than valid trials ($t(21)=-2.373$, $p=.027$,Cohen's $d=0.506$). However $BF_{10}=2.179$ indicates that this is somewhat anecdotal. Further, other measures of gaze cueing with this set of faces in our lab do not replicate this IOR effect, which suggests that it is spurious and thus we do not offer any speculation as to why it may have arisen here.

All other stimuli, timings, design and procedure characteristics are as described in the general methods.

## Results

### *Accuracy (d')*

Crucially, there was a significant main effect of cue validity on WM accuracy, plotted in Figure 4A, $F(2,60)=7.571$, $p=.001$, $\eta p^2=.202$ ($BF_{10}=24.02$, strong evidence). WM was significantly more accurate for squares that had been looked at by the cue face during encoding (valid: M=1.052) compared to squares that were looked away from by the cue face (invalid: M=.912), $t(61)= 3.334$, $p=.004$, Cohen's $d= 0.358$ ($BF_{10}=18.89$, strong evidence). WM was also significantly enhanced for looked at squares compared to those that were encoded in the presence of direct gaze (uncued, direct condition: M=.938), $t(61)= 3.328$, $p=.004$, Cohen's $d= 0.276$ ($BF_{10}=18.583$, strong evidence). No significant difference in WM accuracy was observed between the invalid and direct conditions, $t(61)= -.649$, $p=1$, Cohen's $d= -0.066$ ($BF_{10}=0.17$, moderate evidence for H0).

As expected, the main effect of load was significant, $F(2,60) = 147.640$, $p<.001$, $\eta p^2=.831$, reflecting better performance at load 4 (M = 1.502) than at load 6 (M = .854) ($t=13.062$, $p<.001$, $d= 1.361$) and load 8 (M = .546) ($t=17.010$, $p<.001$, $d=2.177$). Performance at load 6 was also significantly better than at load 8 ($t=6.585$, $p<.001$, $d=.842$). The interaction between load and cue validity was non-significant, $F(4,58)=.865$, $p=.491$, $=\eta p^2=.056$. Hits, FAs, and K capacity estimates for all conditions are reported in Table 1.

*Figure 3 about here*

*Figure 4 about here*

We additionally checked whether the gaze cuing effect was influenced by the gender of the cue face, as well as the location (left or right) of the squares. Neither the main effect of face gender nor the main effect of squares location was significant ($F$ values $\leq .244$, $p$ values $\geq .623$), and neither face gender nor squares location were found to interact with validity ($F$ values $\leq 1.007$, $p$ values $\geq .371$).

The influence of social anxiety and autistic-like traits on gaze cuing effects was also examined. Of the participants included in the main analysis, three (one male) were missing data from the Liebowitz social anxiety scale and two of these same individuals (one male) were also missing data from the ASQ questionnaire and were therefore not included in this analysis. The median social anxiety score was 47, with scores ranging from 4 to 116 (out of maximum of 144). A score of 60 or above demonstrates a socially anxious character (Mennin et al., 2002), thus it is likely that if there was an effect of social anxiety in our study, it would be detectable. A two-tailed bivariate Pearson's correlation was computed using difference scores created by subtracting the d' for the valid minus invalid conditions, to provide an index of cuing magnitude. This was collapsed across load because load was not found to influence cuing effects. We found that social anxiety did not significantly correlate with the magnitude of the gaze cuing effect, $r(59) = -.086$, $p = .515$ ($BF_{10} = 0.2$, moderate evidence for H0). The median ASQ score was 15.5, with scores ranging from 5 to 29 (out of a maximum of 50). The reported control population average score is 16.4 (Baron-Cohen et al., 2001), indicating that there was enough variation in scores for a difference to be detected if it existed. We found that gaze cuing magnitude did not significantly correlate with scores on the ASQ questionnaire, $r(60) = .041$, $p = .754$ ($BF_{10} = 0.17$, moderate evidence for H0). Thus, gaze cuing effects on WM were not significantly modulated by social anxiety or autistic-like traits in the healthy adult population used here.

*Table 1 about here*

*Reaction times*

To check whether there was a speed-accuracy trade-off (i.e., whether accurate responses were made more slowly on valid versus invalid trials), a repeated-measures ANOVA was conducted on RTs from correct response trials only (probe present and absent combined) with gaze direction (valid, invalid, direct,) and load (4, 6, 8) as within-subjects variables. While there was a marginally significant main effect of validity on RTs, $F(2,60) = 3.055$, $p = .055$, $\eta p^2 = .092$ ($BF_{10} = 0.83$, inconclusive evidence), this did not reflect a speed-accuracy trade-off as RTs were significantly faster on valid versus invalid trials ($t(61)= 2.491$, $p=.046$, Cohen's $d= 0.088$, $BF_{10}=2.378$ anecdotal evidence). No significant difference was observed between the valid and direct conditions ($t(61)= -1.333$, $p=.562$, Cohen's $d= -0.047$) or between invalid and direct conditions, ($t(61)= 1.203$, $p=.234$, Cohen's $d= 0.043$). There was no significant interaction between cue validity and load, $F(4,58)=.293$, $p=.881$, $\eta p^2=.020$. As expected, there was a significant main effect of load on response speed, $F(2,60) = 28.088$, $p<.001$, $\eta p^2= .484$, with quicker RTs at load 4 (744ms) than load 6 (775ms), $t(61)=-5.158$, $p<.001$, $d=-0.256$, and also than at load 8 (797ms), $t(61)=-7.516$, $p<.001$, $d=-0.418$. Responses at load 6 were also significantly faster than at load 8, $t(61)= -3.837$, $p<.001$, $d=-0.177$.

Here we provide the first evidence that gaze cues can enhance WM. Enhanced WM for squares in the valid versus direct condition is a particularly important indicator that WM accuracy was boosted in comparison to baseline. Also of note, equivalent WM for squares in the invalid versus direct conditions shows that squares that were looked away from (essentially ignored) by the cue face were not deprioritised by participants compared to baseline.

## Experiment 2: Arrow Cue

While the results of Experiment 1 indicate that gaze cues enhance WM, it remains only an assumption that *'joint attention'* processes were engaged in this task. We cannot definitively conclude that these effects were driven by the social nature of the cue and therefore the mutual focus of two individuals *per se*. Any form of attention shift might contribute to enhanced WM, simply because more attention was allocated to the encoding items. The finding that measures of social anxiety and autistic-like traits did not exert an influence on cuing effects on WM adds further doubt to the social nature of these findings.

Given that non-social, symbolic cues are also shown to reliably orient attention and enhance detection and discrimination of targets in the traditional cuing literature, it is entirely possible that any spatial cue could influence WM in the same manner. Therefore, in Experiment 2 we replicated the WM task but used an arrow cue instead. Arrows are found to be symbolic cues of attention in much the same way as gaze cues (e.g. Ristic et al., 2002), and are therefore often used in order to assess the social versus symbolic nature of gaze cuing. If the effect found in Experiment 1 is simply due to more attention being drawn towards validly cued items, we would expect to find a similar influence of arrow cues on WM in Experiment 2.

## Method

### *Participants*

Fifty nine naive adult volunteers (19 male; mean age 25 years) from the University of Aberdeen completed the study, some in exchange for monetary reimbursement.

### *Stimuli, Design, & Procedure*

The arrow cue was created by superimposing a left arrow symbol (<) and a right arrow symbol (>) in the centre of a horizontal line (109x5 pixels, subtending $3.2^{o}$ x $0.13^{o}$: see Figure 1B). This arrow cue is more symmetrical in nature than other traditional arrow cues (in which the directional arrow element appears at the end of the line), and thus removes the potential confound that asymmetry of the arrow cue serves to orient attention rather than the symbolic meaning. To better mimic the directional shift from direct to averted conveyed by the gaze cues used in Experiment 1, and to provide an equivalent non-cued direct baseline, on each trial the arrow cue began as a vertical line (5x36 pixels, subtending $0.15^{o}$x$0.92^{o}$) bisecting the centre of the horizontal line. On ‘direct’ trials this non-directional image remained on screen, and on valid/invalid cuing trials the vertical line was replaced by either the right or left arrow element (see Figure 1B).

A pilot study (see general methods section) using 22 naïve and independent participants, confirmed that this particular arrow cue reliably oriented attention to a laterally presented asterisk at 500ms SOA (valid: M = 350ms; invalid: M = 368ms; $t$(21)=3.320, $p$=.003, Cohen's $d$=0.708, $BF_{10}$=12.89 - see figure 3). The arrow cue also oriented attention at 150ms SOA (valid: M =395ms; invalid: M = 412ms; $t$(21)=3.740, $p$=.001, Cohen's $d$=0.797; $BF_{10}$=30.37), and had a moderate cuing effect at 1000ms ($t$(21)=-2.072, $p$=.051, Cohen's $d$=0.442; $BF_{10}$=1.326). We did not employ any individual difference measures.

All other stimuli, timings, design and procedure characteristics are as described in the general methods.

## Results

### *Accuracy (d')*

Crucially, there was a non-significant main effect of cue validity on WM performance, $F$(2,57)= 1.402, $p$=.257, $\eta p^2$=.047($BF_{10}$=0.189, Moderate evidence for H0; see Figure 4B). This null effect of the arrow cue on WM was further supported by a priori uncorrected t tests: valid (M=1.059) versus invalid (M=1.009), $t$(58)=0.969, $p$=.337,Cohen's $d$=0.126 ($BF_{10}$=0.222); valid versus direct (M=.986), $t$(58)=1.668, $p$=.101, Cohen's $d$=0.217 ($BF_{10}$=0.524); invalid versus direct, $t$(58)= .624 $p$=.535, Cohen's $d$=0.081 ($BF_{10}$=0.171).

As expected, there was a significant main effect of load, $F$(2,57) =119.465, $p$<.001, $\eta p^2$=.807. Performance at load 4 (M =1.549) was better than at load 6 (M =.882) ($t$(58)=12.392 , $p$<.001, Cohen's $d$=1.13), and load 8 (M = .623) ($t$(58)=15.252, $p$<.001,Cohen's $d$=1.986). Performance at load 6 was also significantly better than at load 8 ($t$(58)=5.364, $p$<.001, Cohen's $d$=.698). The interaction between validity and load was non-significant, $F$(4,55)=.747, $p$=.565, $\eta p^2$=.051. Hits, FAs, and K capacity estimates for all conditions are reported in Table 1.

*Reaction times*

To check whether there was a speed-accuracy trade-off, a repeated measures ANOVA was conducted on RTs from correct trials (probe present and probe absent combined) with cue direction (valid, invalid, direct), and load (4, 6, 8) as within-subject variables. The main effect of cue validity was non-significant, $F(2,57) =2.453$, $p=.095$, $\eta p^2= .079$, and there was a significant interaction between validity and load, $F(4,55) =3.373$, $p=.015$, $\eta p^2= .197$. This was driven by load 6, where the valid items were reacted to significantly faster than the invalid items, $p=.005$. These data help to indicate that there is no speed-accuracy trade-off. There was a significant main effect of load on retrieval response speed, $F(2,57) =28.849$, $p<.001$, $\eta p^2= .503$, with faster RTs at load4 than load 6 ($t(58)=-6.800$, $p<.001$, $d=-0.885$), load 4 than load 8 ($t(58)=-47.371$, $p<.001$, $d=-0.960$) and load 6 than load 8 ($t(58)= -2.931$, $p=.014$, $d=-0.382$).

It is clear from the pilot study that the arrow cue is an effective cue, and, though not significantly, appears somewhat more effective at cuing attention than the gaze cue utilised (see figure 3). It is unclear why an arrow oriented attention a little more strongly than gaze in our pilot study, but it is a highly learned and salient symbol that is solely designed to convey directionality. A face looking in a particular direction on the other hand is not just a directional cue, but can signal more complex social nuances (see discussion). This may be why the arrow influences speed on the simple target location task a little more than gaze, as there is no other element to interpret and process than pure directionality.

Despite the robust ability of the arrow to orient attention, this cue does not effectively modulate WM. In order to further understand the differences between the gaze and arrow cue effects on WM, we compared the difference between the memory for squares seen in the valid vs invalid locations (i.e., cuing magnitude) for both the gaze and arrow conditions. To do this we calculated difference scores by subtracting the d’ memory scores for the invalid condition from the valid scores, averaged across loads. We then compared this value between the gaze and arrow experiments in a one-tailed independent samples t-test. We found that the cuing effect on WM was moderately larger in the gaze study (M=.14, SD =.33), than that found in the arrow study (M=.05, SD=.39), though this did not reach statistical significance $t(119) = 1.372$, $p = .086$, Cohen's $d = .249$ ($BF_{10} = 0.816$; see Figure 5). However, as demonstrated by the Bayesian statistics, which for the valid versus invalid arrow condition is $BF_{10} = 0.22$, the data in Experiment 2 is approximately four times more likely to have occurred under the null hypothesis, than under the experimental hypothesis. This contrasts with Experiment 1 where $BF_{10} = 18.89$, meaning that this gaze cuing effect on WM was almost 19 times more likely to have occurred under the experimental hypothesis. This is remarkable because both cues demonstrate similar effects on basic attention orienting as evidenced in our pilot studies, yet only the gaze cue has a reliable influence on WM. The lack of a significant statistical difference in the size of the cuing effect on WM for gaze versus arrow cues means we cannot definitively conclude that the arrow does not influence WM at all, but what is clear is that gaze is certainly more powerful. We therefore conclude that the influence of eye gaze on WM found in Experiment 1 is more than simple orienting of attention in response to meaningful directional information.

## Experiment 3: Motion Cue

This final experiment was conducted to directly investigate the role of low-level kinetic information in relation to the gaze cuing effect we found in Experiment 1. The use of a dynamic gaze shift in Experiment 1, where the pupil and iris moved from direct to averted, means that it is plausible that simple motion information alone oriented attention to influence WM, and the social nature of the gaze was irrelevant. The arrow cue we used in Experiment 2 perhaps had some element of implied motion (it changed from a vertical line to converging arrow lines on cued trials), making this unlikely to be the case. However, the arrow cue still retained learned, symbolic directional information. Further, both the arrow and gaze cues used were presented at SOAs of 500ms, considered to be within the realms of volitional attentional control (broadly accepted to be from around 200-300ms, Müller & Rabbitt, 1989), and within the period in which cuing effects are most reliably found for both cue types (between 300ms and 700ms SOA, e.g Frischen et al., 2007; Marotta, Lupiáñez, Martella, & Casagrande, 2012). Here we wanted to minimise the contribution of learned, symbolic information and examine low-level motion effects on purely reflexive attention orienting and WM.

To investigate this we used a non-social, non-learned motion cue, which we previously established to orient attention only at 150ms SOA, not 500ms, indicating that this cue triggers purely reflexive shifts of attention (see methods). We designed this cue to comprise a horizontal line bisected by a shorter vertical line, which shifts in space to approximately the same combined degree as a two pupils and irises. In order to understand if the effect found in Experiment 1 was simply due to low-level kinetic cuing, we replicated Experiment 1 using this motion cue.

## Method

### *Participants & Apparatus*

Thirty two naive adult volunteers (6 male; mean age 28 years) from the University of Aberdeen completed the study, some in exchange for course credit, and some for monetary reimbursement. This is notably different to experiment 1 where we tested 64 participants. This initial large number was tested due to the inclusion of individual difference measures, which are not used in experiments 2 and 3. We tested a larger sample for experiment 2 in order to conduct a cross-cue comparison of cues that are known to orient attention at 500ms

SOA. However, here we used the result of experiment one to determine an appropriate sample size to test the effect of low level motion. An a-priori power analysis was conducted using G-power focussing on the main effect of gaze, using partial eta squared ($\eta p^2 = .106$) to estimate sample size. This analysis informed that a sample size of 25 would be required in order to obtain a significant result. Therefore, if our line cue can influence memory in the same way as eye gaze, a sample size of 32 was deemed sufficient.

*Stimuli, Design & Procedure*

The motion cue comprised the central 'direct' cue used in Experiment 2, with leftward and rightward motion created by moving the vertical bisecting line 0.48° to the left or right of the centre point (see Figure 1C). All other stimuli were the same as in Experiment 1. A pilot study employing the traditional target location cuing paradigm (using 22 naïve and independent participants), showed that this motion cue reliably oriented attention to a laterally presented asterisk at 150ms SOA (valid: M =386ms; invalid: M =402ms; $t(21)= 3.103$, $p=.005$, Cohen's $d=0.661$, $BF_{10}=8.377$; see figure 3) but had negligible impact on attention orienting at 500ms SOA (valid: M = 357ms; invalid: M = 360ms; $t(21)=.478$, $p=.638$, Cohen's $d=0.102$, $BF_{10}=0.247$) and at 1000ms SOA (valid: M = 362 ms; invalid: M = 362ms; $t(21)=.043$, $p=.966$, Cohen's $d=0.009$, $BF_{10}=0.223$). Therefore, while in Experiments 1 and 2 we used the known cuing interval of SOA500ms, here we used SOA150ms to ensure that the squares appeared for encoding within the attention window induced by the motion cue.

All other stimuli, timings, design and procedure characteristics are as described in the general methods.

## Results

*Accuracy (d')*

Crucially, there was a non-significant main effect of cue validity on WM performance, *F*(2,30)= .259, *p*=.773, $\eta p^2$=.017 ($BF_{10}$=0.12, moderate evidence for H0) - see Figure 4C. The null effect of the motion cue on WM was further supported by a priori uncorrected t tests: valid (M=.871) versus invalid (M=.916), *t*(31)=-.720, *p*=.477,Cohen's *d*=-.104 ($BF_{10}$=0.240); valid versus direct (M=.903), *t*(31)=-.494, *p*=.625, Cohen's *d*=-.089 ($BF_{10}$=0.211); invalid versus direct, *t*(31)= .211 *p*=.834, Cohen's *d*=.030 ($BF_{10}$=0.193). The Bayes factors demonstrate that the results are at least four times more likely to have occurred under the null hypothesis than the experimental hypothesis.

As expected, there was a significant main effect of load, *F*(2,30) =45.943, *p*<.001, $\eta p^2$=.754. Performance at load 4 (M =1.344) was better than at load 6 (M = .796), *t*=7.632, *p*<.001, Cohen's *d*=1.149, and load 8 (M = .594.), *t*=9.674, *p*<.001,Cohen's *d*=1.818. Performance at load 6 was also significantly better than at load 8, *t*=4.117, *p*=.001, Cohen's *d*=.728. The interaction between validity and load was non-significant, *F*(4,28)=.751, *p*=.566, $\eta p^2$=.097. Hits, FAs, and K capacity estimates for all conditions are reported in Table 1.

*Reaction times*

The main effect of cue validity was non-significant, *F*(2,30) =2.448, *p*=.104, $\eta p^2$= .140, as was the interaction between validity and load, *F*(4,28) =.994, *p*=.427, $\eta p^2$= .124, indicating no speed-accuracy trade-off. There was a significant main effect of load on response speed, *F*(2,30) =11.240, *p*<.001, $\eta p^2$= .428, with faster RTs at load 4 (867ms) than load 6 (904ms), *t*(31)=-4.322, *p*<.001, *d*=-0.177, and also than at load 8 (939ms), *t*(31)=-4.612, *p*<.001, *d*=-0.334. Retrieval RTs at load 6 were also significantly faster than at load 8, *t*(31)= -3.212, *p*=.009, *d*=-0.167.

These results indicate that while simple motion can reliably and reflexively orient attention to locate a target, it does not modulate WM. Therefore, the influence of eye gaze on

WM found in Experiment 1 is more than simple orienting of attention in response to kinetic directional information. It appears to be social in nature and reflect engagement in joint attention.

## Discussion

The aim of the current study was to determine whether joint attention could influence WM for coloured squares. To do this we combined a traditional gaze cuing paradigm with an established visual WM task using coloured squares (Experiment 1). To investigated the social nature of gaze cue effects, and determine whether any results reflected true joint attention versus general attentional effects, we ran the same task using non-social cues. In Experiment 2 we used an arrow to determine the contribution of learned, symbolic cues, and in Experiment 3 we used a low-level, simple motion cue to determine the contribution of kinetically cued directional information. In all experiments, cues were valid, invalid, or remained ‘direct’ (uncued baseline), and were completely non-predictive of where the squares would appear for encoding. Pilot studies affirmed that the gaze, arrow, and motion cues we used here all reliably oriented attention to speed localisation of a simple target on valid versus invalid trials. However, only gaze cues influenced WM for coloured squares (Experiment 1). WM accuracy was better for squares that were looked at by the cue face (validly cued) compared to squares that were looked away from (invalidly cued) and those encoded in the context of direct gaze (uncued). In addition, retrieval was faster for validly cued versus invalidly cued squares. Arrow and motion cues did not influence WM accuracy or retrieval RTs, indicating that the effects of gaze cues on WM are social in nature, cannot be attributed solely to learned symbolism or low-level motion of the pupil and iris, and thus reflect joint attention processes.

There are three possible mechanisms that could underpin gaze-induced enhancement of WM: (1) facilitative encoding of looked at items, (2) inhibition of looked away from items, and (3) a combination of facilitation and inhibition. Our data best fits the first model, as we find no evidence of inhibition: if looked away from items were inhibited, this would be evidenced by poorer WM for items in the invalid versus direct condition. Our data instead show that WM did not differ for items that were invalidly cued compared to those that were encoded in the context of direct gaze, indicating that squares ‘ignored’ by the cue face were not significantly deprioritised relative to squares that were simply not looked at. We do find clear evidence for gaze-induced facilitation, wherein valid gaze cues enhanced WM relative not only to invalid gaze cues but to the uncued (direct gaze) baseline. Thus, jointly attended information appears to have received special priority in WM.

Increased attention is shown to enhance visual perceptual processing (e.g. Carrasco, Ling, & Read, 2004; Pestilli & Carrasco, 2005), and the close connection between attention and WM (Cowan et al., 2005) may suggest that enhanced WM for looked at squares simply reflects general increased attention at encoding. It may therefore seem unremarkable that we obtain a WM benefit for items cued by gaze. However, the clear absence of any reliable cuing effects on WM using two other powerful, non-social cues suggests that the effect is not simply a result of greater attention allocation to the cued location resulting in enhanced perceptual encoding of the colours. It could also be argued that gaze cues enhanced encoding because participants were able to move their eyes within 500ms to foveate the memoranda when it briefly appeared (overt eye movements occur beyond 200-300ms, Müller & Rabbitt, 1989). If this were true then the arrow with the same cue-target interval of 500ms should have similarly enhanced WM, but it did not.

Our findings instead add to the growing evidence that engaging in joint attention serves to do more than simply enhance perceptual processing of objects. As Becchio, Bertone, & Castiello (2008) state, “…objects falling under the gaze of others acquire properties that they would not display if not looked at.” (pp 254). For example, Bayliss, Paul, Cannon, & Tipper, (2006) showed that everyday objects were rated as more likeable when they were encoded under joint attention (compared to invalidly cued and direct conditions), while arrow cues did not influence likeability. This effect of joint attention on object evaluation has been shown in a variety of studies in a variety of ways (Bayliss et al., 2007; Capozzi, Bayliss, Elena, & Becchio, 2015; King, Rowe, & Leonards, 2011; Manera, Elena, Bayliss, & Becchio, 2014; Treinen, Corneille, & Luypaert, 2012; van der Weiden, Veling, & Aarts, 2010). This suggests that when an item is appraised in the context of another's gaze it can increase in affective value, and raises the possibility that jointly attended information is tagged with higher status and therefore prioritised in WM. Further, research by Shteynberg, (2010) points towards a 'social tuning' effect, where stimuli experienced as part of a social group are more prominent both cognitively and behaviourally.

Gaze cues indicate more than just directionality of interest or desirability, they indicate the intent to act upon this interest and thus engage in goal-directed thought and action (e.g. Johansson, Westling, Bäckström, & Flanagan, 2001; Land, Mennie, & Rusted, 1999; Land & Tatler, 2009). Thus, engaging in joint attention not only results in a shared focus of attention but in shared goals, which is important for predicting imminent behaviour and enabling collaboration (Huang, Andrist, Sauppé, & Mutlu, 2015). We tend to remember things that are important to our goals (Altmann & Trafton, 2002; Montagrin, Brosch, & Sander, 2013), and raising the value of information via joint attention may serve to enhance goal-directed WM processes. The arrow and line cues simply provide directional information and do not signal goal-directed intent. This could be a likely explanation of why these non-social cues do not influence WM in our study, but do influence the speed to locate a simple target in a similar manner as gaze cues. The role of intent in gaze cues can also help explain the influence of gaze but not arrow cues on long-term memory (Dodd et al., 2012) and object evaluation (e.g., Bayliss et al., 2006).

It would be pertinent for future research to probe more specifically the stage(s) of WM which are influenced by joint attention - encoding, maintenance, retrieval - and in particular manipulate the duration of the maintenance interval to examine the durability of such potential value tags.

Our finding that WM is influenced by gaze cues but not non-social cues also lends further support to the broader notion that not all attention orienting is the same, and that social attention may be special – at least in some contexts. There is ongoing debate as to whether gaze and arrow cues operate on different attention systems. Many traditional cuing tasks which involve detection, localisation, or discrimination of a simple target find a similar timecourse of cuing effects from gaze and arrow cues, and show that both cues can elicit a mix of reflexive (exogeneous) and volitional (endogenous) orienting responses (e.g. Galfano et al., 2012; Tipples, 2002). However, Hietanen and colleagues (2006) found that, despite showing similar behavioural cuing effects, a different cortical network mediated orienting to gaze cues compared to arrow cues. It may also be possible that the level of processing required in response to the target task may to some degree determine whether gaze cues operate in a unique manner. While there has been no systematic review of gaze versus non-social cuing effects as a function of task, there is growing evidence of gaze-specific effects in tasks which involve higher order processes such as affective judgement (Bayliss et al., 2006), long-term memory (Dodd et al., 2012), and working memory as found here. Future behavioural and neuropsychological investigation is needed to unpack this and provide clarity on how exactly gaze and non-social cues differ and in which contexts.

Although WM load effects were not a key question here, we presented a variety of loads in order to assess whether any effects of cuing were modulated by task difficulty / resource demands. It is possible that the onset of a large number of coloured squares (i.e., load 8) would capture more attention than the onset of fewer squares (i.e., load 4), and potentially override the attentional effects of the cue, thus resulting in weaker cuing at higher versus lower loads. However, we did not find any interaction between cue validity and WM load (regardless of cue type), and further analyses showed that cuing magnitude (valid minus invalid accuracy scored) did not significantly differ between any load condition in all three experiments (all *p*s > .356).

To conclude, our finding that joint attention enhances WM indicates that we not only share the focus of attention with others, but can share more deeply their immediate, short-term goals. Engaging with other individuals and looking where they look provides a window into their intentions towards other objects and people in the immediate environment. We used only simple stimuli - coloured squares - and one key question moving forward is how joint attention may modulate WM for more complex, meaningful, and emotional information.

# References

Abrams, R. A., & Christ, S. E. (2003). Motion onset captures attention. *Psychological Science*, *14*(5), 427–432. doi:10.1111/1467-9280.01458

Altmann, E. M., & Trafton, J. G. (2002). Memory for goals: an activation-based model. *Cognitive Science*, *26*(1), 39–83. doi:10.1207/s15516709cog2601_2

Awh, E., Vogel, E. K., & Oh, S. H. (2006). Interactions between attention and working memory. *Neuroscience*, *139*(1), 201–8. doi:10.1016/j.neuroscience.2005.08.023

Baddeley, A. D. (2007). *Working memory, thought and action*. Oxford: Oxford University Press.

Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., & Clubley, E. (2001). The Autism Spectrum Quotient (AQ): Evidence from Asperger Syndrome/High-Functioning Autism, Males and Females, Scientists and Mathematicians. *Journal of Autism and Developmental Disorders*, *31*(1), 5–17. doi:10.1023/A:1005653411471

Bayliss, A. P., di Pellegrino, G., & Tipper, S. P. (2005). Sex differences in eye gaze and symbolic cueing of attention. *The Quarterly Journal of Experimental Psychology. A, Human Experimental Psychology*, *58*(4), 631–650. doi:10.1080/02724980443000124

Bayliss, A. P., Frischen, A., Fenske, M. J., & Tipper, S. P. (2007). Affective evaluations of objects are influenced by observed gaze direction and emotional expression. *Cognition*, *104*(3), 644–53. doi:10.1016/j.cognition.2006.07.012

Bayliss, A. P., Paul, M. A., Cannon, P. R., & Tipper, S. P. (2006). Gaze cuing and affective judgments of objects: I like what you look at. *Psychonomic Bulletin & Review*, *13*(6), 1061–1066. doi:10.3758/BF03213926

Bayliss, A. P., & Tipper, S. P. (2005). Gaze and arrow cueing of attention reveals individual differences along the autism spectrum as a function of target context. *British Journal of Psychology*, *96*(1), 95–114. doi:10.1348/000712604X15626

Becchio, C., Bertone, C., & Castiello, U. (2008). How the gaze of others influences object processing. *Trends in Cognitive Sciences*, *12*(7), 254–258. doi:10.1016/j.tics.2008.04.005

Capozzi, F., Bayliss, A. P., Elena, M. R., & Becchio, C. (2015). One is not enough: Group size modulates social gaze-induced object desirability effects. *Psychonomic Bulletin & Review*, *22*(3), 850–5. doi:10.3758/s13423-014-0717-z

Carrasco, M., Ling, S., & Read, S. (2004). Attention alters appearance. *Nature Neuroscience*, *7*(3), 308–313. doi:10.1038/nn1194

Charman, T. (2003). Why is joint attention a pivotal skill in autism? *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *358*, 315–24. doi:10.1098/rstb.2002.1199

Charman, T., Baron-Cohen, S., Swettenham, J., Baird, G., Cox, A., & Drew, A. (2000). Testing joint attention, imitation, and play as infancy precursors to language and theory of mind. *Cognitive Development*, *15*(4), 481–498. doi:10.1016/S0885-2014(01)00037-5

Charman, T., Swettenham, J., Baron-Cohen, S., Cox, A., Baird, G., & Drew, A. (1997). Infants with autism: An investigation of empathy, pretend play, joint attention, and imitation. *Developmental Psychology*, *33*(5), 781–789. doi:10.1037/0012-1649.33.5.781

Cleveland, A., Schug, M., & Striano, T. (2007). Joint attention and object learning in 5- and 7- month-old infants. *Infant and Child Development*, *16*(3), 295–306. doi:10.1002/icd.508

Cowan, N. (2000). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, *24*(1), 87–185. doi:10.1017/S0140525X01003922

Cowan, N. (2010). The Magical Mystery Four: How is Working Memory Capacity Limited, and Why? *Current Directions in Psychological Science*, *19*(1), 51–57. doi:10.1177/0963721409359277.The

Cowan, N., Elliott, E. M., Saults, S. J., Morey, C. C., Mattox, S., Hismjatullina, A., & Conway, A. R. a. (2005). On the capacity of attention: Its estimation and its role in working memory and cognitive aptitudes. *Cognitive Psychology*, *51*(1), 42–100. doi:10.1016/j.cogpsych.2004.12.001

Dodd, M. D., Weiss, N., McDonnell, G. P., Sarwal, A., & Kingstone, A. (2012). Gaze cues influence memory…but not for long. *Acta Psychologica*, *141*(2), 270–5. doi:10.1016/j.actpsy.2012.06.003

Doherty, M. (2006). The development of mentalistic gaze understanding. *Infant and Child Development*, *186*, 179–186. doi:10.1002/icd.

Driver, J., Davis, G., Ricciardelli, P., Kidd, P., Maxwell, E., & Baron-Cohen, S. (1999). Gaze Perception Triggers Reflexive Visuospatial Orienting. *Visual Cognition*, *6*(5), 509–540. doi:10.1080/135062899394920

Emery, N. J. (2000). The eyes have it: the neuroethology, function and evolution of social gaze. *Neuroscience & Biobehavioral Reviews*, *24*(6), 581–604. doi:10.1016/S0149-7634(00)00025-7

Farroni, T., Csibra, G., Simion, F., & Johnson, M. H. (2002). Eye contact detection in humans from birth. *Proceedings of the National Academy of Sciences of the United States of America*, *99*(14), 9602–9605. doi:10.1073/pnas.152159999

Fresco, D. M., Coles, M. E., Heimberg, R. G., Liebowitz, M. R., Hami, S., Stein, M. B., & Goetz, D. (2001). The Liebowitz Social Anxiety Scale: a comparison of the psychometric properties of self-report and clinician-administered formats. *Psychological Medicine*, *31*(6), 1025–1035. doi:10.1017/S0033291701004056

Friesen, C. K., & Kingstone, A. (1998). The eyes have it! Reflexive orienting is triggered by nonpredictive gaze. *Psychonomic Bulletin & Review*, *5*(3), 490–495. doi:10.3758/BF03208827

Frischen, A., Bayliss, A. P., & Tipper, S. P. (2007). Gaze cueing of attention: visual attention, social cognition, and individual differences. *Psychological Bulletin*, *133*(4), 694–724. doi:10.1037/0033-2909.133.4.694

Frith, C. D., & Frith, U. (2006). The neural basis of mentalizing. *Neuron*, *50*(4), 531–4. doi:10.1016/j.neuron.2006.05.001

Galfano, G., Dalmaso, M., Marzoli, D., Pavan, G., Coricelli, C., & Castelli, L. (2012). Eye gaze cannot be ignored (but neither can arrows). *The Quarterly Journal of Experimental Psychology*. doi:10.1080/17470218.2012.663765

Hanna, J. E., & Brennan, S. E. (2007). Speakers' eye gaze disambiguates referring expressions early during face-to-face conversation. *Journal of Memory and Language*, *57*, 596–615. doi:10.1016/j.jml.2007.01.008

Hietanen, J. K., Nummenmaa, L., Nyman, M. J., Parkkola, R., & Hämäläinen, H. (2006). Automatic attention orienting by social and symbolic cues activates different neural networks: An fMRI study. *NeuroImage*, *33*(1), 406–413. doi:10.1016/j.neuroimage.2006.06.048

Hommel, B., Pratt, J., Colzato, L., & Godijn, R. (2001). Symbolic control of visual attention. *Psychological Science*, *12*(5), 360–365. doi:10.1111/1467-9280.00367

Hood, B., Willen, J., & Driver, J. (1998). Adult's eyes trigger shifts of visual attention in human infants. *Psychological Science*, *9*(2), 131–134. doi:10.1111/1467-9280.00024

Horley, K., Williams, L. M., Gonsalvez, C., & Gordon, E. (2004). Face to face: visual scanpath evidence for abnormal processing of facial expressions in social phobia. *Psychiatry Research*, *127*(1-2), 43–53. doi:10.1016/j.psychres.2004.02.016

Huang, C.-M., Andrist, S., Sauppé, A., & Mutlu, B. (2015). Using gaze patterns to predict task intent in collaboration. *Frontiers in Psychology*, *6*(July), 1049. doi:10.3389/fpsyg.2015.01049

Johansson, R. S., Westling, G., Bäckström, A., & Flanagan, J. R. (2001). Eye-hand coordination in object manipulation. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *21*(17), 6917–6932.

Kane, M. J., Bleckley, M. K., Conway, A. R. A., & Engle, R. W. (2001). A Controlled-Attention View of Working-Memory Capacity. *Journal of Experimental Psychology: General*, *130*(2), 169–183. doi:10.1037//0096-3445.130.2.169

King, D., Rowe, A., & Leonards, U. (2011). I Trust You; Hence I Like the Things You Look at: Gaze Cueing and Sender Trustworthiness Influence Object Evaluation. *Social Cognition*, *29*(4), 476–485. doi:10.1521/soco.2011.29.4.476

Knoeferle, P., & Kreysa, H. (2012). Can Speaker Gaze Modulate Syntactic Structuring and Thematic Role Assignment during Spoken Sentence Comprehension? *Frontiers in Psychology*, *3*(538). doi:10.3389/fpsyg.2012.00538

Land, M., Mennie, N., & Rusted, J. (1999). The roles of vision and eye movements in the control of activities of daily living. *Perception*, *28*(11), 1311–1328. doi:10.1068/p2935

Land, M., & Tatler, B. (2009). *Looking and Acting: Vision and Eye Movements in Natural Behaviour*. Oxford: Oxford University Press.

Langner, O., Dotsch, R., Bijlstra, G., Wigboldus, D. H. J., Hawk, S. T., & van Knippenberg, A. (2010). Presentation and validation of the Radboud Faces Database. *Cognition & Emotion*, *24*(8), 1377–1388. doi:10.1080/02699930903485076

Leys, C., Ley, C., Klein, O., Bernard, P., & Licata, L. (2013). Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median. *Journal of Experimental Social Psychology*, *49*(4), 764–766. doi:10.1016/j.jesp.2013.03.013

Love, J., Selker, R., Marsman, M., Jamil, T., Dropmann, D., Verhagen, A. J., … Wagenmakers, E. J. (2015). JASP (Version 0.7) [Computer software].

Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, *390*(6657), 279–81. doi:10.1038/36846

Manera, V., Elena, M. R., Bayliss, A. P., & Becchio, C. (2014). When seeing is more than looking: Intentional gaze modulates object desirability. *Emotion*, *14*(4), 824–32. doi:10.1037/a0036258

Marotta, A., Lupiáñez, J., Martella, D., & Casagrande, M. (2012). Eye gaze versus arrows as spatial cues: Two qualitatively different modes of attentional selection. *Journal of Experimental Psychology: Human Perception and Performance*, *38*(2), 326–335. doi:10.1037/a0023959

Mennin, D. S., Fresco, D. M., Heimberg, R. G., Schneier, F. R., Davies, S. O., & Liebowitz, M. R. (2002). Screening for social anxiety disorder in the clinical setting : using the Liebowitz Social Anxiety Scale. *Journal of Anxiety Disorders*, *16*(6), 661–673. doi:10.1016/S0887-6185(02)00134-2

Montagrin, A., Brosch, T., & Sander, D. (2013). Goal conduciveness as a key determinant of memory facilitation. *Emotion*, *13*(4), 622–8. doi:10.1037/a0033066

Moukheiber, A., Rautureau, G., Perez-Diaz, F., Soussignan, R., Dubal, S., Jouvent, R., & Pelissolo, A. (2010). Gaze avoidance in social phobia: objective measure and correlates.

*Behaviour Research and Therapy*, *48*(2), 147–51. doi:10.1016/j.brat.2009.09.012

Müller, H. J., & Rabbitt, P. M. (1989). Reflexive and voluntary orienting of visual attention: time course of activation and resistance to interruption. *Journal of Experimental Psychology. Human Perception and Performance*, *15*, 315–330. doi:10.1037/0096-1523.15.2.315

Pestilli, F., & Carrasco, M. (2005). Attention enhances contrast sensitivity at cued and impairs it at uncued locations. *Vision Research*, *45*(14), 1867–1875. doi:10.1016/j.visres.2005.01.019

Posner, M. I. (1980). Orienting of attention. *The Quarterly Journal of Experimental Psychology*, *32*(1), 3–25. doi:10.1080/00335558008248231

Richardson, D. C., Street, C. N. H., Tan, J. Y. M., Kirkham, N. Z., Hoover, M. A., & Ghane Cavanaugh, A. (2012). Joint perception: gaze and social context. *Frontiers in Human Neuroscience*, *6*(JULY), 194. doi:10.3389/fnhum.2012.00194

Ristic, J., Friesen, C. K., & Kingstone, A. (2002). Are eyes special? It depends on how you look at it. *Psychonomic Bulletin & Review*, *9*(3), 507–513. doi:10.3758/BF03196306

Ristic, J., & Kingstone, A. (2012). A new form of human spatial attention: Automated symbolic orienting. *Visual Cognition*, *20*(3), 244–264. doi:10.1080/13506285.2012.658101

Ristic, J., & Landry, M. (2014). Combining attention: a novel way of conceptualizing the links between attention, sensory processing, and behavior. *Attention, Perception, & Psychophysics*, *77*(1), 36–49. doi:10.3758/s13414-014-0737-9

Ristic, J., Landry, M., & Kingstone, A. (2012). Automated symbolic orienting: The missing link. *Frontiers in Psychology*, *3*(DEC), 1–7. doi:10.3389/fpsyg.2012.00560

Schneider, W., Eschman, A., & Zuccolotto, A. (2002). E-Prime user's guide. *Pittsburgh, PA: Psychology Software Tools.*

Shteynberg, G. (2010). A silent emergence of culture: The social tuning effect. *Journal of Personality and Social Psychology*, *99*(4), 683–689. doi:10.1037/a0019573

Striano, T., Chen, X., Cleveland, A., & Bradshaw, S. (2006). Joint attention social cues influence infant learning. *European Journal of Developmental Psychology*, *3*(3), 289–299. doi:10.1080/17405620600879779

Taylor, T. L., & Klein, R. M. (2000). Visual and motor effects in inhibition of return. *Journal of Experimental Psychology: Human Perception and Performance*, *26*, 1639–1656. doi:10.1037/0096-1523.26.5.1639

Tipples, J. (2002). Eye gaze is not unique: automatic orienting in response to uninformative arrows. *Psychonomic Bulletin & Review*, *9*(2), 314–318. doi:10.3758/BF03196287

Tipples, J. (2008). Orienting to counterpredictive gaze and arrow cues. *Perception & Psychophysics*, *70*(1), 77–87. doi:10.3758/PP.70.1.77

Tomasello, M. (1988). The role of joint attentional processes in early language development. *Language Sciences*, *10*(1), 69–88. doi:10.1016/0388-0001(88)90006-X

Tomasello, M., & Farrar, M. J. (2015). Joint Attention and Early Language, *57*(6), 1454–1463.

Treinen, E., Corneille, O., & Luypaert, G. (2012). L-eye to me: The combined role of Need for Cognition and facial trustworthiness in mimetic desires. *Cognition*, *122*(2), 247–251. doi:10.1016/j.cognition.2011.10.006

Tummeltshammer, K. S., Wu, R., Sobel, D. M., & Kirkham, N. Z. (2014). Infants track the reliability of potential informants. *Psychological Science*, *25*(9), 1730–1738. doi:10.1177/0956797614540178

van der Weiden, A., Veling, H., & Aarts, H. (2010). When observing gaze shifts of others enhances object desirability. *Emotion*, *10*(6), 939–943. doi:10.1037/a0020501

Wu, R., Gopnik, A., Richardson, D. C., & Kirkham, N. Z. (2011). Infants learn about objects from statistics and people. *Developmental Psychology*, *47*(5), 1220–1229. doi:10.1037/a0024023

Wu, R., & Kirkham, N. Z. (2010). No two cues are alike: Depth of learning during infancy is dependent on what orients attention. *Journal of Experimental Child Psychology*, *107*(2), 118–136. doi:10.1016/j.jecp.2010.04.014

Xu, S., Zhang, S., & Geng, H. (2011). Gaze-induced joint attention persists under high perceptual load and does not depend on awareness. *Vision Research*, *51*(18), 2048–56. doi:10.1016/j.visres.2011.07.023

# Tables and Figures

**Figure 1**. Illustration of the trial procedure and SOA conditions for experiments 1 (Panel A), 2 (Panel B), and 3 (Panel C). In Experiments 1 and 2 cue-target SOA was 500ms, while in experiment 3 cue-target SOA was 150ms, found to be the optimum timings for facilitative attention shifts. All show the valid, probe present condition.

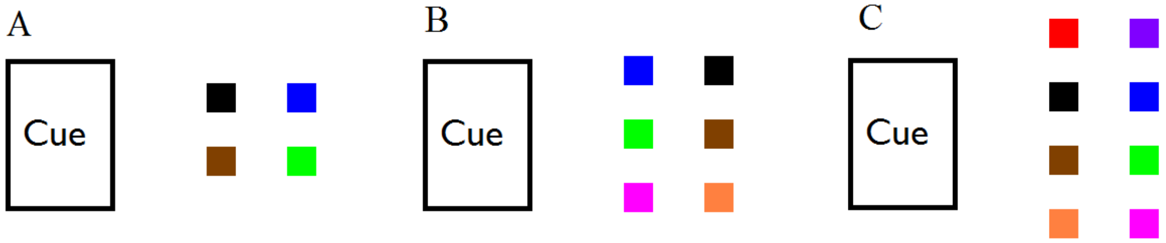

**Figure 2**. Illustration (to scale) of the squares configuration, load 4 (A), 6 (B), and 8 (C). Note that the cue appears in the centre of the screen. The inner square was presented 1.7º visual angle (59 pixels) from the edge of the cue. At load 4, the squares array spanned

2.7°x2.7°; at load 6, 2.7°x3.9°; at load 8, 2.7°x5.8°. The distance between each square was 1.1°.

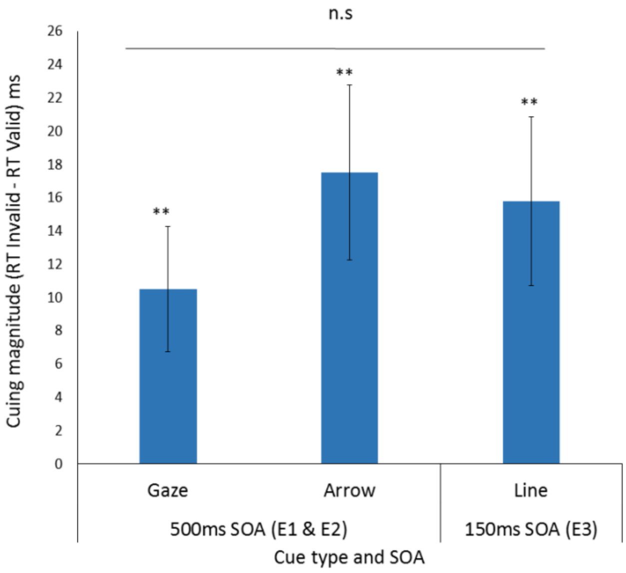

**Figure 3.** Results from the three pilot studies using a traditional target location task. The bars represent the magnitude of the cuing effect. This is calculated by subtracting the reaction times for the valid trials from the reaction times for the invalid trials. Thus a positive value indicates that participants were faster on valid than invalid trials. All cue types resulted in significantly faster response times for validly cued than invalidly cued targets. Error bars represent 1 standard error above and below the mean. Note that the cuing magnitudes across cue types were not significantly different from each other.

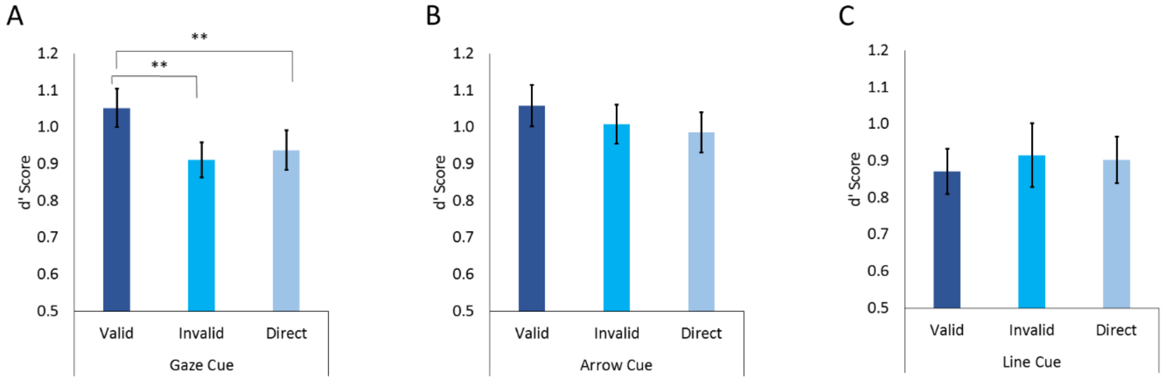

**Figure 4.** Results from Experiments 1 (Gaze cue; Panel A), 2(Arrow cue; Panel B), and 3 (Motion cue; Panel C). WM performance scores using d' are plotted as a function of cue validity. Bars represent 1 standard error above and below the mean.

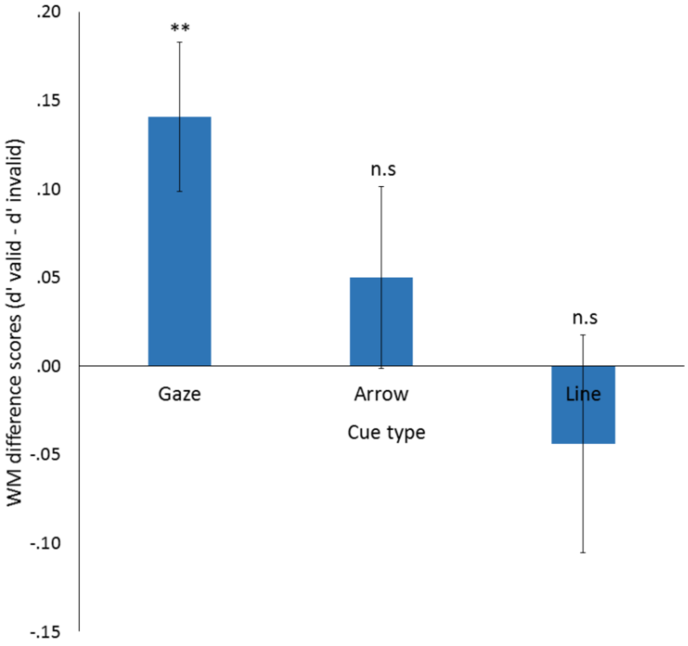

**Figure 5.** WM difference scores for the three experiments. The bars represent the magnitude of the cuing effect on memory. This is calculated by subtracting the d’ score for the invalid trials from the d’ score for the valid trials. Thus a positive value indicates that participants were more accurate on valid than invalid trials. Error bars represent 1 standard error above and below the mean.

**Table 1**: Mean and standard error (in parenthesis) for hits (%), False alarms (%), d', K capacity estimates (Load * (Hits-FAs)), and reaction times (RT; ms) for each cue condition in each experiment at each working memory load (4, 6, 8).

| | **Cue Type** | | **Gaze** | | | **Arrow** | | | **Motion** | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Data** | **Condition** | **Valid** | **Invalid** | **Direct** | **Valid** | **Invalid** | **Direct** | **Valid** | **Invalid** | **Direct** |
| **Hits (%)** | **4** | 73.7 (1.6) | 69.8 (1.6) | 70.8 (1.7) | 71.5 (1.7) | 72.1 (1.6) | 70.3 (1.5) | 67.5 (2.5) | 67.8 (2.5) | 66.7 (1.9) |
| | **6** | 62.1(1.5) | 58.9 (1.9) | 59.3 (2.0) | 59.6 (1.7) | 59.0 (2.1) | 60.1 (2.0) | 56.7 (2.5) | 57.4 (3.1) | 58.7 (2.6) |
| | **8** | 55.5(2.0) | 56.9 (1.7) | 58.0 (1.9) | 59.9 (2.1) | 60.2 (2.3) | 59.9 (2.2) | 56.9 (2.9) | 57.3 (2.6) | 55.8(2.7) |
| | **Mean** | 63.8 (1.4) | 61.9 (1.4) | 62.7 (1.6) | 63.7 (1.5) | 63.8 (1.6) | 63.4 (1.6) | 60.4 (2.1) | 60.9 (2.2) | 60.4(1.7) |
| **FA (%)** | **4** | 19.9 (1.4) | 21.0 (1.4) | 22.6 (1.4) | 19.9 (1.5) | 19.7 (1.3) | 20.5 (1.5) | 23.7 (2.3) | 22.7 (2.2) | 21.5 (2.7) |
| | **6** | 28.6 (1.7) | 32.1 (1.5) | 29.7 (1.6) | 26.1 (1.8) | 31.2 (2.0) | 30.8 (1.9) | 29.8 (2.6) | 32.3 (3.4) | 30.0 (2.6) |
| | **8** | 35.2 (1.8) | 38.2 (1.7) | 38.4 (1.9) | 38.6 (2.5) | 38.2 (2.0) | 38.3 (2.1) | 36.9 (2.7) | 35.2 (2.8) | 39.1 (2.8) |
| | **Mean** | 27.9 (1.4) | 30.4 (1.1) | 30.3(1.3) | 28.2 (1.6) | 29.7 (1.5) | 29.9 (1.5) | 30.1 (2.1) | 30.0 (2.4) | 30.2 (2.4) |
| **d'** | **4** | 1.623 (.082) | 1.455 (.075) | 1.429 (.083) | 1.577 (.086) | 1.584 (.082) | 1.486 (.077) | 1.283 (.099) | 1.363 (.130) | 1.386 (.113) |
| | **6** | 0.960 (.058) | 0.767 (.069) | 0.835 (.073) | 0.972 (.067) | 0.829 (.078) | 0.844 (.073) | 0.779 (.074) | 0.763 (.108) | 0.848 (.088) |
| | **8** | 0.574 (.063) | 0.513 (.047) | 0.551 (.055) | 0.628 (.068) | 0.615 (.054) | 0.626 (.058) | 0.552 (.077) | 0.620 (.075) | 0.475 (.058) |
| | **Mean** | 1.052 (0.052) | 0.912 (0.048) | 0.938 (0.054) | 1.059 (.056) | 1.009 (.053) | 0.986 (.055) | 0.871 (.061) | 0.916 (.087) | 0.903 (.064) |
| **K** | **4** | 2.151 (.091) | 1.953 (.087) | 1.926 (.090) | 2.064 (.094) | 2.094 (.083) | 1.989 (.089) | 1.753 (.122) | 1.806 (.147) | 1.807 (.124) |
| | **6** | 2.008 (.111) | 1.610 (.130) | 1.777 (.149) | 2.011 (.134) | 1.669 (.147) | 1.759 (.144) | 1.613 (.153) | 1.511 (.199) | 1.721 (.163) |
| | **8** | 1.628 (.168) | 1.493 (.132) | 1.564 (.158) | 1.710 (.183) | 1.760 (.155) | 1.727 (.161) | 1.600 (.221) | 1.773 (.213) | 1.335 (.162) |
| | **Mean** | 1.929 (.095) | 1.685 (.086) | 1.755 (.102) | 1.928 (.106) | 1.841 (.101) | 1.825 (.102) | 1.655 (.126) | 1.697 (.154) | 1.621 (.105) |
| **RT (ms)** | **4** | 742 (16) | 748 (16) | 743 (15) | 809 (27) | 798 (25) | 812 (29) | 867 (37) | 868 (37) | 867 (37) |
| | **6** | 765 (15) | 782 (16) | 777 (16) | 826 (27) | 859 (32) | 846 (28) | 899 (37) | 909 (36) | 903 (36) |
| | **8** | 792 (18) | 802 (18) | 797 (17) | 850 (28) | 858 (28) | 864 (33) | 923 (39) | 942 (39) | 952 (41) |
| | **Mean** | 767 (16) | 778 (16) | 772 (15) | 829 (26) | 838 (27) | 841 (29) | 896 (36) | 906 (36) | 907 (37) |

## Author Contributions

M.C. Jackson and S.E.A. Gregory both contributed to the study design. Testing, data collection, and data analyses were performed by S.E.A. Gregory under the supervision of M.C. Jackson. S.E.A. Gregory drafted the manuscript and M.C. Jackson provided critical revisions. Both authors approved the final version of the manuscript for submission.