Flexible egocentricity: Asymmetric switch costs on a perspective-taking task.

Samuel, Steven[a,b]

Roehr-Brackin, Karen[a]

Jelbert, Sarah[b]

Clayton, Nicola S.[b]

[a]University of Essex, UK

[b]University of Cambridge, UK

Word count: 4817 words, excluding title page, footnotes, acknowledgments, tables, figure captions, declarations, and abstract.

Keywords: Perspective-Taking, Egocentric Bias, Director Task.

Address for correspondence: Steven Samuel. University of Cambridge, Department of Psychology, Downing Street, CB2 3EB. Tel: 01223 333563. Email: ss2391@cam.ac.uk

Abstract

We gave 40 participants a task in which they needed to select target objects from an array according to the instructions of either an informed director (who shared their perspective of the array) or an ignorant director (whose view of the array was restricted owing to barriers). Importantly, sometimes only one of the directors was visible, and on some trials when both directors were present participants were required to switch between perspectives. We found that participants were faster to select items from the informed director's perspective than the ignorant director's perspective, but that they slowed when there was a visible but inactive second director. Crucially, relative to non-switch trials where the same perspective was taken twice consecutively, participants exhibited a significant cost of switching between perspectives when returning to take their own perspective, but not when switching to the other point of view. We interpret these results as evidence that participants inhibit their more salient perspective in order to adopt another's, and then incur an asymmetric switch cost as a result. This suggests that although we are egocentric by default, our egocentricity is effectively, albeit temporarily, eliminated if we have just adopted an alternative frame of reference.

Humans are commonly thought to be biased towards their own egocentric view of the world. We over-ascribe to others our own beliefs (Ross, Greene & House, 1977), knowledge (Birch & Bloom, 2007), and visual perspective (Apperly, Carroll, Samson, Humphreys, Qureshi, & Moffitt, 2010; Epley, Morewedge, & Keysar, 2004; Keysar, Lin, & Barr, 2003). When we take visual perspectives, we experience more interference from our own viewpoint when taking another's than the other way around (Ferguson, Apperly, & Cane, 2017; Samson, Apperly, Braithwaite, Andrews, & Bodley–Scott, 2010). However, studies such as these and others (e.g. Tversky & Hard, 2009) also show that despite this apparent primacy of egocentrism, we can be sensitive enough to others' visual perspectives to process them even when there is no overt reason to do so. Additionally, some researchers have found that we integrate multiple strands of information, such as what is visible to both parties or has already been mentioned in speech (i.e. in common ground), suggesting that the egocentrism is thus perhaps not so fundamental a characteristic of our perspective-taking behaviour (Brown-Schmidt, Gunlogsen, & Tanenhaus, 2008; Brown-Schmidt & Hanna, 2011; Yoon & Brown-Schmidt, 2014). The present study was designed to inform this debate by testing the limits of our egocentricity with respect to visual perspective-taking.

The Director Task (Keysar, Barr, Balin, & Brauner, 2000; Keysar et al., 2003) has become one of the most common tests of perspective-taking in recent years (Dumontheil, Apperly, & Blakemore, 2010; Dumontheil, Hillebrandt, Apperly, & Blakemore, 2012; Dumontheil, Küster, Apperly, & Blakemore, 2010; Cane, Ferguson, & Apperly, 2017; Legg, Olivier, Samuel, Lurz, & Clayton, 2017). In the task, a participant is instructed to select and move items on a series of shelves (a 'grid') which are placed between the participant and an 'ignorant' (in non-computerized versions, typically a confederate) director. Due to a number of occlusions in the grid, the ignorant director has a restricted view of the array of objects, and hence an instruction from the ignorant director to select an object in the grid is ambiguous if the instruction describes two items—one the ignorant director *can* see and one he cannot. Figure 1 (panel A) illustrates an example using an image from the present experiment, where the male avatar on the far side of the grid is the ignorant instructor. The image also includes an 'informed' director (female avatar) who shares the participant's view of the array and hence allows for instructions to be given that do not require the participant to take an alternative perspective (e.g. Apperly et al., 2010). The ignorant director's view therefore corresponds to the 'other' perspective, and the informed director's view corresponds to the self-perspective. Depending on which director gives the instruction to select 'the top clock', the target is either the clock in the second row from the top (according to the ignorant director's point of view) or the one in the top left corner (according to the informed director's view). Across both 'real-life' and computerized versions of the Director Task, the typical finding is that participants make more errors and/or are slower when selecting targets from a perspective that is not their own, or when selecting a target for which there is a competitor in privileged ground (i.e. that only the participant can see). These outcomes are consistent with the view that the egocentric perspective is privileged (e.g. Apperly et al., 2010; Epley, Keysar, Van Boven, & Gilovich, 2004; Keysar et al., 2000; Keysar et al., 2003; Legg et al., 2017; Samuel, Roehr-Brackin, & Roberson, 2016).

But what are the limits of this apparent privilege? Is it possible to regulate egocentricity in visual perspective-taking? The questions we ask in the present study are i) do *expectations* regarding whether perspectives might need to be alternated modulate our egocentricity?; and ii) do we achieve such modulation by flexibly inhibiting our egocentricity in real-time such that we may appear less egocentric, or cease to be egocentric at all, in the immediate moments after we have been required to ignore our own point of view?

In order to test our first hypothesis, we compared performance on trials from the informed director when the ignorant instructor was visible versus when the informed director was alone. If participants are slower to take their own perspective in the presence of an ignorant director, then we will have obtained evidence that egocentricity can be modulated by expectations, specifically the expectation that the need to take the alternative viewpoint may be imminent. This is a particularly conservative means of testing our hypothesis because the participant could theoretically simply ignore the ignorant director on any given trial, as their perspective was irrelevant to the successful completion of the informed director's instruction. To make this test more conservative still, we also did not tell participants that directors would sometimes be absent. Thus, we relied on the participants' ability to perceive for themselves which directors were present or not. To ensure that participants were not simply adopting a more cautious approach when two directors were present than when there was only one, perhaps in anticipation of the need to switch, we also manipulated the presence or absence of the *informed* director. This would create a baseline measure of the simple interference effect of a second visible director.

To test our second hypothesis, we also gave participants grids where both directors were on-screen and both also gave instructions. Importantly, each grid incorporated multiple instructions, and hence multiple trials. On 'switching grids' participants heard instructions from both directors at unpredictable times. If egocentricity can be temporarily inhibited to take a non-egocentric perspective, then we should expect participants to find it harder to switch *back* to informed director's (previously inhibited) perspective than switch *to* the ignorant director's perspective. This is because we incur an asymmetric switch cost when returning to a more salient rather than less salient frame of reference; the former requires greater inhibition to overcome, and the subsequent lifting of this inhibition incurs a processing cost. For example, we are slower to switch back to speaking in our more salient first language after speaking in our second than vice-versa (Meuter & Allport, 1999). We are also slower to respond to a congruent trial on Stroop- and Simon-like tasks if it comes after an incongruent trial than after another congruent trial (e.g. Egner, 2007). We hypothesized that if the egocentric perspective takes primacy, then performance should be faster on non-switch trials from this perspective than from the 'other' perspective, but on switch trials this advantage should be reduced.

Method

Participants

We recruited 40[1] native-English-speaking participants (M_{age} = 21yrs, *SD* = 4, eight males, 31 females, one undisclosed/non-binary). Ethical approval was obtained from the University of Cambridge Psychology Research Ethics committee. All participants gave informed consent before participating.

Director Task

[1] This figure conformed to a power analysis based on a pilot study that had revealed a significant difference between the presence or absence of an irrelevant director. The details of this pilot are available from the first author.

Each grid consisted of a 4 x 4 set of shelves with nine objects, three each of identical clocks, cups and vases (see Figures 1A-C). Participants heard three-word instructions (through headphones) all limited to the type: "The [top/bottom] [clock/cup/vase]". One director (henceforth *informed* director) stood on the near side of the grid, and shared the participant's ('self') perspective. Another director (henceforth *ignorant* director) stood on the other side of the grid ('other' perspective) and had a limited view of the objects owing to three occlusions. As a result of this limited view, the participants needed to pay attention to the perspective of the director. This was indexed by the gender of the director's voice. For example, in the case of Figure 1 (panel A), the top cup, top clock or bottom vase would describe different targets depending on whom the instruction came from. Participants responded with mouse clicks on the chosen target. Director gender was fully counterbalanced between participants (i.e. for half the participants the ignorant director was always female).

Different from other Director Tasks, grid types varied according to the number of directors who were present on screen (one or two) and the number of directors who actually gave verbal instructions for that grid. On '*no-switching*' grids only one director gave instructions, but that director could be alone (one director condition) or accompanied (two director condition). When only one director was visible, the other was replaced by a lampstand in order to maintain a physical presence in that spot (see Figures 1B and 1C). On *switching* grids, both directors were visible, and crucially both gave instructions. This format resulted in 'switch' trials, where the director who gave the instruction could be different from the director who spoke on the previous trial, and non-switch trials, where the instruction came from the same director as the previous trial. The results of switching and non-switching grids were analysed separately.

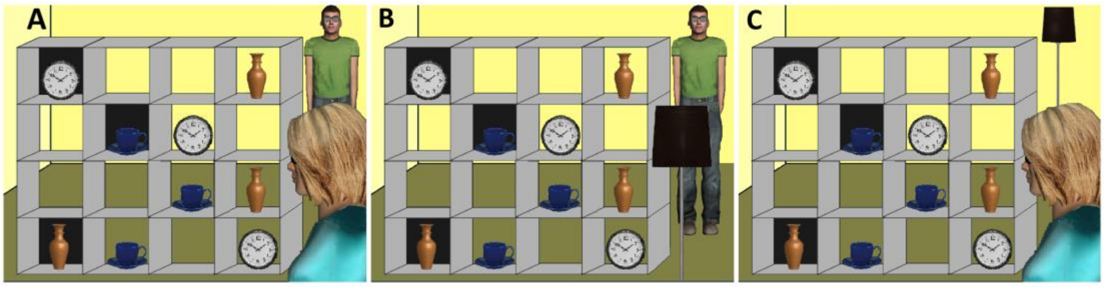

Figure 1. Examples of the visual stimuli used in the experiment. A) two-director grid– this could be either a No-switching on Switching grid, depending on whether just one director gave instructions or both did on that grid; B) one-director grid (ignorant director only) C) one-director grid (informed director only). These stimuli depict examples with the ignorant director played by a male; half of the participants received a version with the ignorant director played by the female.

A total of 216 trials were distributed over 42 grids of 4-6 trials with a fixed 3500ms inter-trial interval. The grids and directors for that grid were presented on screen 5000ms prior to the first instruction. *Experimental* trials concerned those occasions where the instruction would refer to a different object if the non-target perspective was erroneously selected, and *filler* trials involved a target that would be the same regardless of perspective. Across the 28 no-switching grids, there were 18 each of experimental and filler trials from each director (informed and ignorant) in each of the one-director and two-director conditions, making a total of 72 non-switch trials and 72 filler trials. Apart from the presence or absence of the second director, all the visual and audio stimuli were constant across the one-director and two-director conditions. Crucially, each director was on screen for the same length of time across the task as a whole.

There were also 36 experimental trials and a matched number of filler trials across the 14 switching grids. Half of the experimental trials were switch trials, so there were 9 non-switch and 9 switch trials from each director. These grids were interspersed among the non-switching grids, and ensured participants could not always maintain the perspective of the first director they heard on any given grid. To reinforce this, the practice grid that all participants performed prior to the task proper was also a switching condition grid. Participants only moved on the task proper once they performed this grid error-free. To avoid any potential order effects, we randomly generated two grid sequences and created two more by reversing the order of grids on them. Coupled with the counterbalancing of director gender, this meant that there were eight versions of the experiment, each performed by five participants. The layout of each grid array and full details of the sequencing of each version of the experiment are included in the supplemental materials (SOM1 and SOM2 respectively).

Participants were told to take the perspective of the director who spoke into account when clicking on objects in the grid, and were told to respond as quickly and as accurately as possible. They were *not* told that a director would be replaced on some grids by a lampshade, or that only one director would give instructions on certain grids. The task took around 20 minutes to complete.

Results

No-switching grids

Accuracy. We first analysed those grids where only one director gave instructions, but another director may or may not have been present. By definition, all trials on these grids were non-switch trials. A total of 34 experimental trial errors were time-outs (<1% experimental trials), and these were included as errors in the accuracy data, but were removed from the response time analyses. Responses made 800ms or faster after trial onset (i.e. before noun onset) almost always occurred after a timeout trial, indicating that they were late responses to the previous trial, and these 28 trials (< 1%) were removed entirely as a result. The remaining errors were clicks on other objects or areas of the screen. These were included in accuracy analyses as errors, but excluded from response time analyses. Mean accuracy was very high, and is listed in Table 1. The distribution of accuracy rates was not normal. Friedman's two-way analyses of variance by ranks found no variation between accuracy on these four trial types (Z = 0.277, p = .964).

Table 1. Accuracy (SD in parentheses) on no-switching grids (experimental trials).

No-Switching grids	Self perspective		Other perspective	
One director	97% (5%)		98% (4%)	
Two directors	97% (6%)		97% (5%)	
Switching grids	Non-Switch	Switch	Non-Switch	Switch
Two directors	96% (6%)	93% (9%)	97% (11%)	93% (9%)

Response times. Mean response times were calculated from noun onset, which itself was calculated for each individual instruction/audio file. As the data were normally distributed in each cell (Shapiro-Wilks tests, all ps >.5), we proceeded with parametric tests.

We conducted a 2: Perspective (Self vs. Other) x 2: Presence (One Director vs. Two Directors) fully within-subjects ANOVA. The condition means are displayed in Figure 2. The analysis revealed a main effect of Presence (F(1,39) = 26.622, MSE = 5093, p < .001, η_p^2 = .406), with participants 58ms slower when both directors were present than when only one was visible (SE = 11ms, 95% CI = [35,81]), and a main effect of Perspective (F(1,39) = 19.117, MSE = 16926, p = <.001, η_p^2 = .329), with participants 90ms slower to take the other, ignorant director's perspective (SE = 21ms, 95% CI = [48,132]). There was no evidence of an interaction (F(1,39) = 0.133, MSE = 4781, p = .717, η_p^2 = .003).

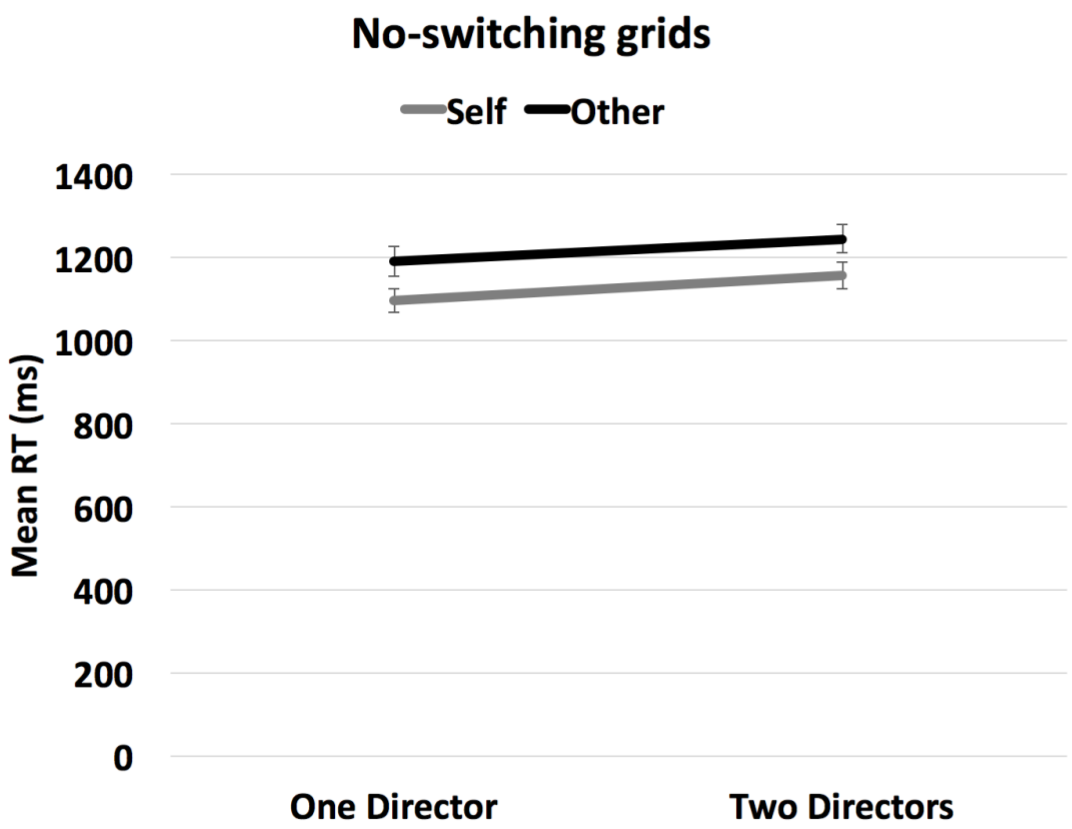

Figure 2. Mean response times on No-switching grids.

Switching grids

Accuracy. We next analysed the data for those 12 grids where both directors were present *and* each gave instructions. A total of 26 (<2%) experimental trial errors were time-outs, and responses made 800ms or faster after trial onset (i.e. before noun onset) were removed entirely (17 trials, <2%). Mean accuracy is displayed in Table 1.

Friedman's two-way analyses of variance by ranks found significant variation in accuracy between the four trial types ($Z = 15.732$, $p = .001$), but follow-up pairwise comparisons applying the Bonferroni correction found no differences to be significant (adjusted *ps* > .09).

Response times. As the data were normally distributed in each cell (Shapiro-Wilks tests, all *p*s >.3), we proceeded with parametric tests.

We conducted a 2: Perspective (Self vs. Other) x 2: Switch (Non-Switch vs. Switch) fully within-subjects ANOVA. The condition means are displayed in Figure 3. The analysis revealed main effects of Perspective ($F(1,39) = 24.611$, $MSE = 11770$, $p < .001$, $\eta_p^2 = .387$), with participants 85ms faster (SE = 17ms, 95% CI [50,120] to take the informed (self) than ignorant (other) director's perspective, and also a main effect of Switch ($F(1,39) = 31.510$, $MSE = 6951$, $p < .001$, $\eta_p^2 = .447$), with participants 74ms slower to respond on switch trials than non-switch trials (SE = 13ms, 95% CI [47, 101]). Crucially, there was also a significant interaction ($F(1,39) = 35.917$, $MSE = 11003$, $p < .001$, $\eta_p^2 = .479$). We examined this interaction using follow-up post-hoc pairwise comparisons with the Bonferroni correction. On non-switch trials, participants were 184ms faster from the informed director's (self) perspective than from the ignorant director's (other) perspective (SE = 23ms, 95% CI [137, 232], adjusted $p < .001$). On switch trials however, participants performed at similar speeds across both perspectives ($M_{Diff} = 14$ms, SE = 24ms, 95% CI [-35, 64], adjusted $p = .1$). Additionally, on self-perspective trials participants were 173ms slower on switch than non-switch trials (SE = 21ms, 95% CI [130, 217], adjusted $p < .001$), but on trials from the ignorant director's (other) perspective no such difference was found ($M_{Diff} = 25$ms, SE = 21ms, 95% CI [-17, 68], adjusted $p = .47$). In sum, switching back to the easier self-perspective incurred a cost roughly equivalent to the advantage of holding the self-perspective in the first place.

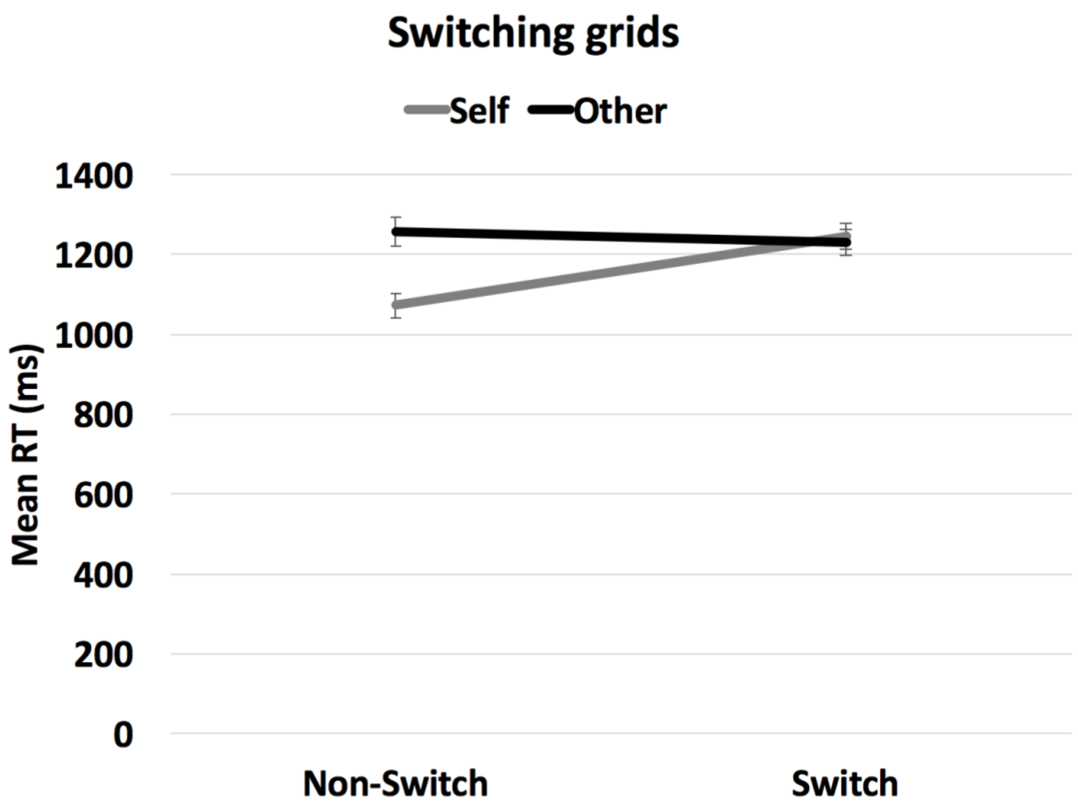

Figure 3. Mean response times on grids where both directors were active (switching grids).

Discussion

We explored whether i) participants' egocentricity would be modulated by the mere presence of an ignorant director, and ii) whether participants showed evidence of inhibiting their own perspective in order to take another's. The results pointed to four main findings. Firstly, participants were faster overall on trials from the self-perspective than the other perspective. Secondly, participants were slower when switching perspectives than when holding them. The third finding concerned our first hypothesis; we found that participants were slower to respond when there were two directors on screen than when there was only one. However, this cost was not restricted to the appearance of a redundant ignorant director, but was equally present when there was a redundant informed director, ruling out an account by which participants modulated their egocentricity specifically. The fourth finding concerned our second hypothesis; we found that the advantage conferred by taking one's own egocentric perspective disappeared when switching perspectives. In other words, the primacy of egocentricity was eliminated on trials immediately following trials from the ignorant director. We discuss these four findings in turn.

Firstly, the finding that participants were faster to take the informed director's perspective than the ignorant director's when either was alone was entirely expected, fitting as it does the theory that we have an egocentric bias in visual perspective-taking (e.g. Apperly et al., 2010; Epley et al., 2004; Keysar et al., 2003). We feel that there are at least three specific reasons for why participants were faster from the informed than the ignorant director's perspective on one-director trials: i) The informed director's perspective could be easier because it omits the need to process the more complex spatial structure that occlusions create; ii) the target of an instruction from the informed director was always the best perceptual match for the description (i.e., the top vase is factually the top vase; e.g., Hanna, Tanenhaus & Trueswell, 2003; see also Heller, Parisien & Stevenson, 2016); and iii) the informed director's *visual* perspective is reinforced with the extra activation that comes from its being shared with the participants'. We do not prefer one of these potential accounts over another, and of course more than one may be involved at the same time.

Our second finding, namely that participants were slower on switch trials than non-switch trials, was also expected. Participants do typically perform more slowly when they are required to switch between visual perspectives within blocks of trials (e.g. Ferguson et al., 2017; Samson et al., 2010; though see Samuel et al., 2016). Explanations that have been put forward for this include evidence from eye tracking data that participants conduct a broader visual search on switch trials, or possibly experience greater interference between self and other states (e.g. Ferguson et al., 2017).

The same processes that make switching perspectives harder likely account for the slower response times when two directors are on screen than only one. Interestingly, we found that the size of this effect was similar regardless of whether it was the ignorant or the informed director that was visible but inactive on that grid. The present study therefore establishes that this interference effect occurs not only in the presence of a director who takes a different view from the participant's, but even extends to a director who *shares* the participant's egocentric perspective. Given that the informed director was in effect carrying a 'redundant' perspective because it was the participant's own in any case, this result points strongly to an account by which the presence of a second director resulted in the expectation that a more taxing switch trial may be imminent. This creates a rather unusual 'egocentric' interference effect caused not by one's own knowledge, but by an avatar who shares it. This has important ramifications for perspective-taking research that uses proxies such as avatars to represent self-perspective trials in a task; it appears that it cannot be taken for granted that just because an avatar and participant share a perspective that the participant performs the same with or without the proxy.

Our most interesting finding was that the advantage of the self-perspective on non-switch trials disappeared on switch trials. This was driven by a statistically significant slowdown on switch trials to the informed director. Switching to the ignorant director did not cause any such slowdown. This pattern of results appears to be consistent with the suggestion that the egocentric perspective was the more salient by default, and hence required a disproportionate degree of inhibition relative to the other perspective, a finding that is in turn consistent with the evidence from other tasks that report so-called asymmetrical switch costs that we described earlier (e.g. Egner, 2007; Meuter & Allport, 1999; see Koch, Gade, Shuch, & Philipp, 2010, for a discussion of potential mechanisms underpinning asymmetric switch costs). It is also consistent with the broader view that domain-general, top-down processes can modulate our ability to take non-egocentric perspectives (e.g. Brown-Schmidt, 2009; Cane et al., 2017; Lin et al., 2010). For example, Lin et al. (2010) found that participants with greater working memory resources needed less time between first fixating and then selecting targets in a Director Task, and Cane et al. found that rewarding participants led them to make greater use of perspective information (i.e. what was in privileged or common ground). In the original computerized version of the Director Task, Apperly and colleagues (2010) also suggested that general executive processes were implicated in the ability to successfully take the ignorant director's perspective. However, studies also exist that suggest top-down processes do *not* influence our ability to take perspective (e.g. Brown-Schmidt, 2012; Ryskin, Brown-Schmidt, Canseco-Gonzalez, Yiu, & Nguyen, 2014). It may be that more local, task-based factors influence any relationship, such as whether we take the perspective of a person or an avatar (see for example Skaratt, Cole, & Kuhn, 2012). Our finding also seems to contradict that of Bradford, Welsh, and Gomez (2015), who reported an asymmetric switch cost favouring more efficient switches *to* the self-perspective. This inconsistency might be explained by differences in task aims and design. For example, in Bradford and colleagues' study the other person was introduced suddenly at the moment of the switch and not before, which may have increased processing demands relative to switches to the other perspective in our study, and relative to switches to a self perspective (which is instead present throughout by default). Perhaps more importantly, our study was concerned with testing visual perspective-taking, whereas Bradford et al. were investigating belief attributions. It may be that different processes are engaged in switching between visual perspectives and switching between belief states.

Overall, however, we feel that the results of this study appear not to offer unqualified support to one theory of perspective-taking or another. On the one hand, the results appear to favour the view that the egocentric perspective is the most salient by default, firstly because trials from this perspective were performed more quickly, and secondly because we do not see a switch cost on trials from the ignorant director. However, it is also clear that this egocentricity is flexible; it can be temporarily eliminated by the mere act of taking the ‘other’ frame of reference as our starting point and then having to adjust back to our own. In this sense, our results might also be viewed as supporting accounts that posit a less pivotal role of egocentrism in such tasks (e.g. Brown-Schmidt & Hanna, 2011). Important questions for future research thus concern at what point the balance of perspective salience might tip in favour of another’s viewpoint for a more sustained period; how many times do we need to take another person’s perspective before we are required instead to suppress that perspective? How quickly might such a reversal of salience take to decay? And given the fluid and efficient perspective-taking in linguistic interaction (Brown-Schmidt & Hanna, 2011), is egocentricity more flexible in language-based tasks?

In sum, we found that although we are believed to be inherently egocentric, taking an alternative point of view is enough to temporarily suspend the salience of our own perspective. As such, the primacy of our egocentricity may be more flexible, or fragile, than we think.

Acknowledgments

This work was supported by an ESRC grant (ES/M008460/1) (SS & NSC; PI = NSC) and European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013)/ERC Grant Agreement No. 3399933 (SJ & NSC) awarded to NSC. With thanks to Ian Apperly for use of original stimuli, and Edward Legg for comments on a draft of this manuscript.

Disclosure of interest

The authors report no conflicts of interest.

References

Apperly, I. A., Carroll, D. J., Samson, D., Humphreys, G. W., Qureshi, A., & Moffitt, G. (2010). Why are there limits on theory of mind use? Evidence from adults' ability to follow instructions from an ignorant speaker. *The Quarterly Journal of Experimental Psychology*, *63*(6), 1201–1217.

Birch, S. A., & Bloom, P. (2007). The curse of knowledge in reasoning about false beliefs. *Psychological Science*, *18*(5), 382–386.

Bradford, E. E., Jentzsch, I., & Gomez, J. C. (2015). From self to social cognition: theory of mind mechanisms and their relation to executive functioning. *Cognition*, *138*, 21-34.

Brown-Schmidt, S. (2009). The role of executive function in perspective taking during online language comprehension. *Psychonomic bulletin & review*, *16*(5), 893-900.

Brown-Schmidt, S. (2012). Beyond common and privileged: Gradient representations of common ground in real-time language use. *Language and Cognitive Processes*, *27*(1), 62-89.

Brown-Schmidt, S., Gunlogson, C., & Tanenhaus, M. K. (2008). Addressees distinguish shared from private information when interpreting questions during interactive conversation. *Cognition*, *107*(3), 1122-1134.

Brown-Schmidt, S., & Hanna, J. E. (2011). Talking in another person's shoes: Incremental perspective-taking in language processing. *Dialogue & Discourse*, *2*(1), 11-33.

Cane, J. E., Ferguson, H. J., & Apperly, I. A. (2017). Using perspective to resolve reference: The impact of cognitive load and motivation. *Journal of Experimental Psychology: Learning, Memory & Cognition (online first)*.

Dumontheil, I., Apperly, I. A., & Blakemore, S. J. (2010). Online usage of theory of mind continues to develop in late adolescence. *Developmental science*, *13*(2), 331-338.

Dumontheil, I., Hillebrandt, H., Apperly, I. A., & Blakemore, S. J. (2012). Developmental differences in the control of action selection by social information. *Journal of Cognitive Neuroscience*, *24*(10), 2080-2095.

Dumontheil, I., Küster, O., Apperly, I. A., & Blakemore, S. J. (2010). Taking perspective into account in a communicative task. *Neuroimage*, *52*(4), 1574-1583.

Egner, T. (2007). Congruency sequence effects and cognitive control. *Cognitive, Affective, & Behavioral Neuroscience*, *7*(4), 380-390.

Epley, N., Keysar, B., Van Boven, L., & Gilovich, T. (2004). Perspective taking as egocentric anchoring and adjustment. *Journal of personality and social psychology*, *87*(3), 327-339.

Epley, N., Morewedge, C. K., & Keysar, B. (2004). Perspective taking in children and adults: Equivalent egocentrism but differential correction. *Journal of Experimental Social Psychology*, *40*(6), 760-768.

Ferguson, H. J., Apperly, I., & Cane, J. E. (2017). Eye tracking reveals the cost of switching between self and other perspectives in a visual perspective-taking task. *The Quarterly Journal of Experimental Psychology*, *70*(8), 1646-1660.

Hanna, J. E., Tanenhaus, M. K., & Trueswell, J. C. (2003). The effects of common ground and perspective on domains of referential interpretation. *Journal of Memory and Language*, *49*(1), 43-61.

Heller, D., Parisien, C., & Stevenson, S. (2016). Perspective-taking behavior as the probabilistic weighing of multiple domains. *Cognition*, *149*, 104-120.

Keysar, B., Barr, D. J., Balin, J. A., & Brauner, J. S. (2000). Taking perspective in conversation: The role of mutual knowledge in comprehension. *Psychological Science*, *11*(1), 32-38.

Keysar, B., Lin, S., & Barr, D. J. (2003). Limits on theory of mind use in adults. *Cognition*, *89*(1), 25–41.

Koch, I., Gade, M., Schuch, S., & Philipp, A. M. (2010). The role of inhibition in task switching: A review. *Psychonomic bulletin & review*, *17*(1), 1-14.

Legg, E., Olivier, L., Samuel, S., Lurz, R., & Clayton, N. (2017). Error rate on the director's task is influenced by the need to take another's perspective but not the type of perspective. *Royal Society Open Science*, 4(8), 170284.

Lin, S., Keysar, B., & Epley, N. (2010). Reflexively mindblind: Using theory of mind to interpret behavior requires effortful attention. *Journal of Experimental Social Psychology*, *46*(3), 551-556.

Meuter, R. F., & Allport, A. (1999). Bilingual language switching in naming: Asymmetrical costs of language selection. *Journal of memory and language*, *40*(1), 25-40.

Ross, L., Greene, D., & House, P. (1977). The “false consensus effect”: An egocentric bias in social perception and attribution processes. *Journal of Experimental Social Psychology*, *13*(3), 279–301.

Ryskin, R. A., Brown-Schmidt, S., Canseco-Gonzalez, E., Yiu, L. K., & Nguyen, E. T. (2014). Visuospatial perspective-taking in conversation and the role of bilingual experience. *Journal of Memory and Language*, *74*, 46-76.

Samson, D., Apperly, I. A., Braithwaite, J. J., Andrews, B. J., & Bodley Scott, S. E. (2010). Seeing it their way: evidence for rapid and involuntary computation of what other people see. *Journal of Experimental Psychology: Human Perception and Performance*, *36*(5), 1255-1266.

Samuel, S., Roehr-Brackin, K., & Roberson, D. (2016). ‘She says, he says’: Does the sex of an instructor interact with the grammatical gender of targets in a perspective-taking task? *International Journal of Bilingualism*, *20*(1), 40-61.

Skarratt, P. A., Cole, G. G., & Kuhn, G. (2012). Visual cognition during real social interaction. *Frontiers in human neuroscience*, *6*.

Tversky, B., & Hard, B. M. (2009). Embodied and disembodied cognition: Spatial perspective-taking. *Cognition*, *110*(1), 124-129.

Yoon, S. O., & Brown-Schmidt, S. (2014). Adjusting conceptual pacts in three-party conversation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*(4), 919-937.