CrossMark

RESEARCH REPORT

Gesturing during mental problem solving reduces eye movements, especially for individuals with lower visual working memory capacity

Wim T. J. L. Pouw[1,2] · Myrto-Foteini Mavilidi[2] · Tamara van Gog[1,3] · Fred Paas[1,2]

Abstract Non-communicative hand gestures have been found to benefit problem-solving performance. These gestures seem to compensate for limited internal cognitive capacities, such as visual working memory capacity. Yet, it is not clear how gestures might perform this cognitive function. One hypothesis is that gesturing is a means to spatially index mental simulations, thereby reducing the need for visually projecting the mental simulation onto the visual presentation of the task. If that hypothesis is correct, less eye movements should be made when participants gesture during problem solving than when they do not gesture. We therefore used mobile eye tracking to investigate the effect of co-thought gesturing and visual working memory capacity on eye movements during mental solving of the Tower of Hanoi problem. Results revealed that gesturing indeed reduced the number of eye movements (lower saccade counts), especially for participants with a relatively lower visual working memory capacity. Subsequent problem-solving performance was not affected by having (not) gestured during the mental solving phase. The current findings suggest that our understanding of gestures in problem solving could be improved by taking into account eye movements during gesturing.

Handling editor: Katsumi Watanabe, Tokyo University

Reviewers: Mingyuan Chu, University of Aberdeen, Massimiliano Cappuccio, United Arab Emirates University

✉ Wim T. J. L. Pouw
  pouw@fsw.eur.nl

1 Department of Psychology, Education and Child Studies, Erasmus University Rotterdam, Rotterdam, The Netherlands

2 Early Start Research Institute, University of Wollongong, Wollongong, Australia

3 Department of Education, Utrecht University, Utrecht, The Netherlands

Introduction

Gesturing can benefit problem solving, especially under conditions of high cognitive load (e.g. Chu and Kita 2011; Marstaller and Burianová 2013; for a review see Pouw et al. 2014). Yet, the exact mechanisms through which the cognitive system exploits manual activity are still not clear. Based on the literature discussed below, we hypothesized that gestures (pointing) allow for spatially indexing mental simulations in space, which come to stand in for eye movements that are performed to visually project mental simulations onto the presentation of the task in the external environment (Cappuccio et al. 2013; Cooperrider et al. 2015; Pouw et al. 2014). To test that hypothesis, we investigated whether gesturing (pointing) versus not gesturing during mental problem solving (Tower of Hanoi; hereon TOH) affected eye movements.

Gesturing during problem solving (or mentally problem solving) has been shown to benefit (subsequent) problem-solving performance, especially when cognitive load is high (for a review Pouw et al. 2014). That is, when the task is more complex (e.g. Chu and Kita 2011; Delgado et al. 2011; Logan et al. 2014) and/or when cognitive resources (such as working memory) are limited (e.g. Marstaller and Burianová 2013). For example, participants who spontaneously used pointing gestures, or had been instructed to

Springer

gesture while mentally solving the TOH for 150 s (in silence), subsequently performed better on solving the problem as compared to participants who did not gesture (Pouw et al. under review). However, gesturing was only beneficial for performance compared to not gesturing under conditions of higher cognitive load: for participants with lower visual working memory capacities, and only on more complex trials. In line with these results, there is evidence that gestures are indeed spontaneously employed to compensate for visual processing load: spontaneous gestures have been found to increase in rate when subjects are wearing glasses that project visually complex information compared to when simple information is projected (Smithson and Nicoladis 2014).

The effect of cognitive load on gestures' effectiveness can be interpreted from an embodied and embedded cognition perspective (Cappuccio et al. 2013; Clark 2013; Pouw et al. 2014). According to this interpretation, gestures offer the cognitive system stable extra-neural tools for visuo-spatial thinking from which new or improved cognitive resources can emerge. That is, gestures embed, support and extend ongoing internal cognitive processing (e.g. working memory).

Yet, a major challenge for current research on the role of gesture in problem solving is to specify *how* gestures support cognitive processes (Cappuccio et al. 2013; Cooperrider et al. 2015; Pouw et al. 2014). One potential mechanism was proposed by Cappuccio et al. (2013). Focusing on the role of pointing gestures, they suggest that gesturing during problem solving provides a compensatory mechanism for visual processing: "pointing hence represents a stand-in for the corresponding series of acts of ocular redirection; the benefits received from monitoring these acts affect capabilities such as keeping track of what has been counted, individuating objects, focusing on a particular object, anchoring number words to objects… …double-check, re-organize, concentrate, and parse in time/space the task…" (p. 141).

Indeed, there is evidence that eye movements ("ocular redirection") reflect and even support mental simulations during online problem solving. For example, Kirsh (2009) confronted participants with a tic-tac-toe game, in which they had to keep track of their own and the opponent's moves in mental imagery. It was found that participants (especially those with a low spatial ability) performed better on the most difficult task when they could play the game while looking at a tic-tac-toe matrix as opposed to an empty sheet of paper. The tic-tac-toe matrix allowed a way to "project" mentally simulated information on a presentation of the task in the environment. In similar vein, findings from eye-tracking research on solving the TOH suggest that problem solvers actively explore possible

moves visually when presented with a 2D presentation of the task, anticipating (or simulating) the placement of the disc from one peg to another with an eye movement (e.g. Patsenko and Altmann 2010). As argued by Spivey and Dale (2011), eye-tracking research in problem solving (e.g. Thomas and Lleras 2007) suggests that eye movements not only reflect but also support ongoing problem solving by anchoring cognitive processes in the environment. This visual projection strategy, though, produces substantial cognitive load, because of the need to not only visually plan, but also visually monitor the "correctness" of each step of the mental simulation, mapped onto an external visual presentation that has not (yet) changed. Thus, this strategy might be especially difficult for those with lower visual working memory capacity.

Although not explicitly stated by Cappuccio et al. (2013), gestures are "monitored" through proprioception (i.e. the sense of the relative positions of body and limbs in space; see Pouw et al. 2014 for a discussion). Gesturing, we would suggest therefore, provides an additional non-visual-based spatial presentation that can anchor mental simulations. That gesture's function is (at least in part) proprioceptive, is in line with recent research that shows that gestures affect problem solving of the TOH even when gesturers cannot see their own hands (Cooperrider et al. 2015). When gestures are proprioceptively monitored, it can be hypothesized that gestures can come to "stand-in" for eye movements as an anchor for mental simulations in the external environment, thereby reducing the number of eye movements being made. Furthermore, this effect should be stronger under conditions of higher visual working memory load, that is, when tasks are more complex or (when task complexity is equal) for those individuals who have lower visual working memory capacity.

We investigated this hypothesis in the present study. Participants performed two trials of the TOH of similar complexity: each trial consisted of a 4-disc problem but with normal or inversed rules, wherein each solution path is exactly the same (see method for details). In one of the two trials, participants were instructed to gesture (pointing in silence) during a 60-s mental solving phase that preceded actual problem solving, and in the other trial, participants did not gesture. If pointing gestures indeed allow for spatially indexing a mentally simulated move of a disc in space surrounding the body (peri-personal space), then the need to project information visually onto the 2d presentation of the task becomes functionally redundant, and a lower saccade count would be expected on the gesture trial than on the non-gesture trial. Moreover, we would predict that the function of gesturing is especially relevant to (and therefore exploited by) those with lower WM capacity, as

those with higher WM capacity may be able to easily project mental simulations using a visual strategy. If this prediction is correct, it could provide a functional explanation to why gestures seem especially effective for those with a lower visual working memory capacity (e.g. Marstaller and Burianová 2013; Pouw et al. under review).

Method

Participants and design

This study was approved by the Human Research Committee of the University of Wollongong. A total of 20 adults participated in the present study (employees of the Early Start Institute Wollongong), who were unaware of the hypotheses of the study ($M_{age} = 34.40$, SD = 8.63, age range 24–50 years; 5 males).

A within-subject experimental design was used, meaning that all participants performed two versions of the 4-disc TOH task. Depending on counterbalancing condition, participants were instructed not to gesture or to use pointing gestures during the first or second mental solving phase. Whether they first solved the normal TOH and then the inverted TOH or vice versa, was also counterbalanced between subjects. Each physical solving phase was preceded by a mental solving phase of the task for 60 s.

Before the start of the experiment, participants reported previous experience with the TOH (yes or no) and one participant reported that he had experience with solving the TOH in the past. We did not exclude this participant, because our within-subject design should control for possible confounds of skill in relation to the manipulation. Note, however, that excluding this participant resulted in the same pattern of findings reported in the result section.

Apparatus and materials

Eye-tracking equipment

Eye movements were recorded with SMI eye-tracking glasses 2.0 connected via USB to a smartphone from which the data could be uploaded afterwards. Data were analyzed with SMI BeGaze software (version 3.3). The sampling rate was set at 60 Hz and was bi-ocular. For each participant, before the start of the experiment, a 3-point triangular calibration was performed (distance between participants' eyes and points on the screen: point 1: 175 cm, point 2: 175 cm, point 3: 154 cm; distance between point 1 and point 2: 98 cm, distance between points 1 and 2 with point 3: 56 cm). To verify the accuracy of the calibration, subjects were asked to look at the same points again.

Video screen

All tasks were performed on the computer that projected onto a large LED TV screen, size 167×95 cm. The distance between the eyes and the screen was 165 cm.

Visual patterns test

The visual patterns test (VPT; Della Sala et al. 1997) was a mouse-based task and served as a proxy for visual working memory capacity. We used an adapted version of the VPT (as adapted from and kindly provided by Chu et al. 2013). Participants were shown a matrix, in various patterns, wherein half of the cells (i.e. squares of 14 cm $\times$ 14 cm) were coloured black. Each pattern was displayed for 3 s, after which all the squares turned white. Participants needed to recreate the pattern of black squares by selecting the squares in a non-specific order, which upon selecting would turn black. The VPT consisted of 25 trials, with blocks of 5 trials per difficulty level (from 7 to 11 black squares). Before the start of the task, participants were provided with 2 practice trials (3 and 4 black squares, respectively). If participants failed to recall all the black squares during a given trial, it was scored as an incorrect response. After five consecutive incorrect responses within one difficult block of trials, the experimenter stopped the task. Performance scores were the proportion of correct responses out of all trials.

Tower of Hanoi The TOH was programmed in Adobe Flash and consisted of three evenly spaced pegs (distance between pegs: 41 cm, bases: 29.5×2.5 cm, peg: 2×3.4 cm) with four discs (disc 1: 29×4 cm, disc 2: 24×4, disc 3: 17×4 cm, disc 4: 12×4 cm). In the starting position, all discs were stacked on the outer left peg. In the normal rule TOH, the discs decreased in size (i.e. disc 1–4), and the inverted rule TOH increased in size (i.e. disc 4–1). Discs could be placed on the other pegs during the problem-solving process with the click-and-drag mouse function. The goal of the TOH is to transfer the discs from the left peg to the right peg in the same stacking order, subject to the following rules: (1) only one disc at a time can be moved to another peg, (2) a disc can only be moved if it is on the top of the stack, and (3) only smaller discs can be placed on top of bigger discs (normal TOH set-up) or only bigger discs can be placed on top of smaller discs (inverted TOH set-up).

Procedure

Prior to the experiment, participants provided their written consent. Participants were tested individually with the two experimenters present in the room (but they could not see the

experimenters during the tasks). They were first presented with the VPT. Participants were instructed on the nature of the task and performed two practice trials before the start of the VPT proper. The VPT task took approximately 5 min to complete, and there were no time restrictions for this task.

Subsequently, participants put on the eye-tracking glasses and the eye tracker was calibrated. After successful calibration, a practice TOH task with two discs was presented to participants, and consistent with the counterbalance order, this was a normal TOH or inverted TOH practice task. The experimenter explained the rules of the task (with the third rule depending on assigned condition), and participants then solved the two-disc TOH trial as practice (for both normal and inverted TOH). After each instruction, the experimenter verified whether subjects understood the instructions based on whether they solved the practice trial and participants were also asked to verbally repeat the rules.

After the practice trial, participants were informed that before actually solving a similar 4-disc TOH trial, they would be presented with the begin state of the 4-disc TOH trial (i.e. discs placed on the outermost left peg) and that they should mentally plan the moves in silence for 60 s so they could solve the task as fast as possible directly afterwards. Participants were told that they should rehearse the solving moves repeatedly during this phase. Depending on the counterbalancing condition, participants were instructed to think with their hands using pointing gestures during this mental planning phase in a way that suited them (gesture condition). During this instruction, the experimenter made several pointing gestures directed at the TOH as a cue how gestures could be performed. Participants were additionally instructed that they should not gesture directly in front of their face (this was done to ensure that field of vision was not, or only peripherally occluded by gesturing). In the no gesture condition, participants were asked not to move their hands during the 60 s of mental solving. Directly after the mental solving phases, participants solved the respective 4-disc TOH.

This cycle (practice task, mental solving, actual solving) was repeated twice. Participants either received the normal task first and the inverted second or vice versa (i.e. counterbalanced between participants) and were instructed either to gesture on the first task and not on the second or vice versa (i.e. counterbalanced between participants). Once participants correctly solved the first problem, they automatically proceeded to the next cycle. When participants were unable to solve the task, they automatically proceeded to the next cycle after 5 min. Participants were recorded during the TOH (mental) solving phases with a video camera for the purpose of counting their gestures after the experiment. Finally, participants were debriefed and thanked for their participation.

Scoring and data analysis

Gesture Participants' video data per task were coded for gesture frequency (for an example see Fig. 1). Due to camera malfunction, we could not count gestures of two participants. Gestures were defined as any hand movement of one or both hands from one still point to the next, indicating the travel of a disc from one peg to another (see Garber and Goldin-Meadow 2002). All participants used index-pointing gestures. The first two authors independently counted the gestures, and interrater reliability was high, Pearson's $r = .89$, $p < .001$.

Fig. 1 Example of gesturing during the mental solving phase (1 s per frame). To show where participants look at during gesturing, the last frame is an example of the static Tower of Hanoi presented for 60 s during mental problem solving (inverted rules condition)

Eye movement data The number of saccades within the 60-s mental solving phase per task was generated using default settings of the eye-tracking software SMI BeGaze software (version 3.3) for the exact period of 60 s.

Performance For the two problem-solving trials, we obtained solving speed and number of solving steps (number of mistakes were not counted by the programme). Lower number of solving steps and faster solving speeds reflect a higher performance. For each TOH problem-solving trial, the minimal amount of steps necessary to solve the task were fifteen steps. As the given period of solving a trial was set at 300 s, participants who did not solve the task in 300 s were not scored on performance.

Results

Three participants had to be excluded due to technical issues with the eye-tracking glasses. This resulted in a total sample of 17 participants (M_{age} = 35.24, SD = 9.10, age range 24–50 years; 4 males), wherein counterbalancing resulted in $N = 4$ for gesture-normal set-up, $N = 5$ for no gesture-normal set-up, $N = 4$ for gesture-inverted set-up and $N = 4$ for no gesture-inverted set-up during the first TOH trial (counterbalanced for the second TOH trial). Where TOH performance effects are concerned, an additional 2 participants were not included in the sample as they were not able to solve one of the two TOH trials within 300 s.

Table 1 presents the means, standard deviations and correlations of VPT score, solving steps and solving speed during the solving phase, as well as saccade counts during the mental solving phase. Note that higher VPT scores were associated with fewer fixations and saccades overall ($ps < .034$). Interestingly, however, when partialling out the correlations per condition (gesture vs. no gesture), we found that this overall significant correlation was primarily carried by the no gesture condition (VPT and saccade count: $r = -.541$, $p = .025$). In the gesture condition, there was no significant correlation of VPT with saccade

Table 1 Overall means and standard deviations, and correlations between VPT score, solving time TOH, solving steps TOH and saccade count

	M (SD)	1.	2.	3.
1. VPT score	76 (.13)			
2. Solving speed	89.86 (38.97)	−.054		
3. Solving steps	29.37 (12.02)	.195	.828**	
4. Saccade count	138.08 (25.41)	−.517*	.131	.024

* $p < .05$; ** $p < .01$

count ($r = -.022$, $p = .933$). Note, however, that these correlations did not significantly differ, $p = 0.123$ (see Lee and Preacher 2013). These results suggest that visual working memory capacity was more predictive for saccade count in the no gesture condition.

The mean pointing-gesture frequency (which could only be obtained for 15 participants because two had to be excluded due to camera malfunction) during the mental solving phase was 31.87 (SD = 13.11; minimum gesture frequency = 14, maximum = 57). We found no significant correlations between gesture frequency and VPT score, $r = .13$, $p = .638$. Also, the gesture frequency on the task was not significantly correlated with solving speed on the respective trial (which was preceded by gesturing during the mental solving phase), $r = -.33$, $p = .224$, nor was this the case for solving steps, $r = -.28$, $p = .320$. We also checked whether gesture frequency was associated with saccade and fixation count, but no significant associations were found, saccade count $r = -.02$, $p = .953$, fixation count $r = -.07$, $p = .802$.

Eye movements

To test our main hypothesis whether gesturing leads to lower saccade counts during the mental solving phase as compared to not gesturing, and whether this effect was moderated by visual working memory capacity, we performed two separate mixed-effects analyses of covariance (ANCOVAs) on the number of saccades. For each DV, we examined the within-subject effect of gesturing versus not gesturing, with VPT score as a covariate. We first checked for between-subject effects of counterbalancing order of gesture first versus no gesture first, as well as the order of TOH type (normal vs. inverted) by adding these as between-subject factors, in which there were not: gesture counterbalance order, $F(1, 12) = .22$, $p = .644$, TOH type counterbalance order, $F(1, 12) = .03$, $p = .865$, and interaction, $F(1, 12) = 1.63$, $p = .226$.

The results did reveal a significant relationship on the number of saccades when participants gestured compared to when they did not gesture, $F(1, 12) = 8.34$, $p = .014$, $\eta_p^2 = .41$. Overall, fewer saccades were observed when participants gestured (estimated means saccade count = 124.06, SD = 26.52, 95 % CI 108.36–140.00) than when they did not gesture (estimated means saccade count = 132.88, SD = 39.24, 95 % CI 115.34–151.79) when controlling for the covariate VPT. Moreover, there was a significant interaction of gesture condition and the VPT regarding the number of saccades, $F(1, 12) = 7.32$, $p = .019$, $\eta_p^2 = .38$. In Fig. 2 we have plotted the effect of VPT score on the observed differences of saccade count across gesture condition. As Fig. 2 shows, the reduction in saccades when gesturing compared to not
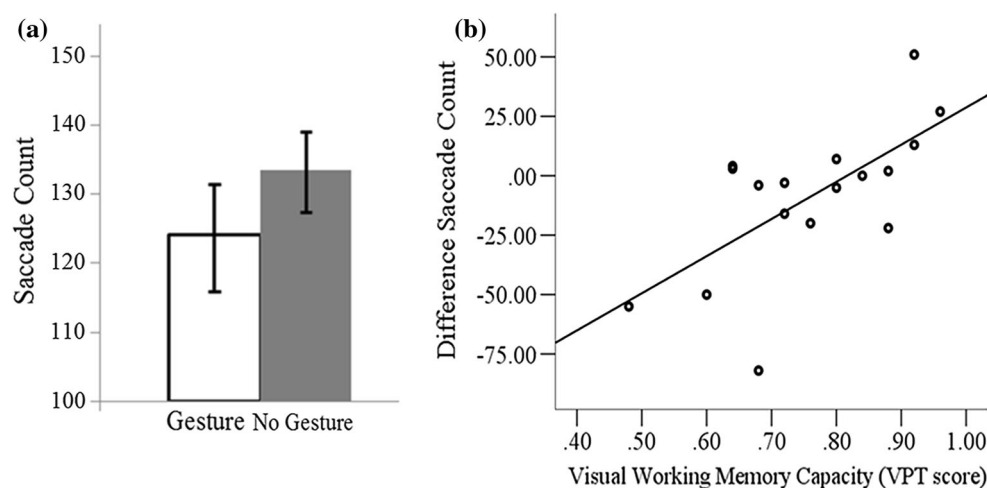
(a)

(b)

Fig. 2 a, b On the *left*, the estimated marginal means and standard errors of the ANCOVA for number of saccades during the 60 s are presented. On the *right*, the difference scores are presented in relation to visual working memory capacity. *Note* On the *right plot*, a negative difference means that lower saccade counts were observed when gesturing was stronger for participants who scored lower on the VPT.[1, 2]

participants gestured versus did not gesture during the mental solving phase. The trend shows that participants with a lower visual working memory capacity were more extremely affected by gesturing, such that a lower saccade count was observed when participants gestured as compared to when they did not gesture

TOH performance

For exploratory purposes, we assessed whether gesture condition and VPT affected performance of the TOH, using two repeated-measures ANCOVAs with solving time or solving steps as the dependent variable, gesture versus no gesture during the mental solving phase as within-subject factor, counterbalancing variables gesture order and TOH type as between-subject variables and the scores on the VPT task as the covariate.

Solving time

No effects of between-subject (i.e. counterbalance) factors gesture order, $F(1, 10) = 1.86$, $p = .202$, TOH order, $F(1, 10) = 0.16$, $p = .699$, or their interaction, $F(1, 10) = 1.97$, $p = .191$, were found on solving time. Furthermore, solving time was not affected by whether participants gestured ($M = 94.24$, SD $= 55.80$, 95 % CI 65.48–124.48) or did not gesture ($M = 85.48$, SD $= 42.03$, 95 % CI 64.55–109.96) during the mental solving phase, $F(1, 10) = 1.25$, $p = .289$. Also, VPT was not significantly co-varying with observed differences, $F(1, 10) = 1.37$, $p = .346$.

Solving steps

No effects of between-subject (i.e. counterbalance) factors gesture order, $F(1, 10) = 0.44$, $p = .523$, TOH order, TOH, $F(1, 10) = 0.99$, $p = .341$, or their interaction, $F(1, 10) = 0.21$, $p = .655$, were found on the number steps taken to solve the problem. Additionally, solving steps were not affected by whether participants gestured ($M = 30.63$, SD $= 14.68$, 95 % CI 22.27–39.01) or did not gesture ($M = 28.13$, SD $= 16.73$, 95 % CI 17.97–38.28) during the mental solving phase, $F(1, 10) = 0.54$, $p = .479$. Also, VPT was not significantly co-varying with observed differences, $F(1, 11) = 0.65$, $p = .437$.

---

[1] Note that since saccade and fixation frequency closely covary, very similar results are obtained when taking into account fixation frequency. A similar repeated-measures analysis of covariance (ANCOVA) was performed with fixation count as the dependent variable. There were no significant between-subject effects of gesture counterbalance order, $F(1, 12) = .05$, $p = .831$, or TOH counterbalance order, $F(1, 12) = .21$, $p = .653$, nor did the interaction of counterbalancing conditions have an effect, $F(1, 12) = 2.48$, $p = .141$. Results revealed significantly lower fixation counts when participants gestured (estimated means $= 143.70$, $SE = 5.38$, 95 % CI 131.97 – 155.44) compared to when they did not gesture (estimated means $= 153.298$, $SE = 6.30$, 95 % CI 139.56 – 167.02), $F(1, 12) = 8.29$, $p = .014$, $\eta_p^2 = .41$. Also, there was a significant interaction between the number of fixations and gesture and the VPT, $F(1, 12) = 7.22$, $p = .020$, $\eta_p^2 = .38$.

[2] 2. As was to be expected given the fixed time available for mental problem solving and the lower fixation count, average fixation duration when gesturing was somewhat higher than when not gesturing, but a similar repeated-measure analysis of covariance (ANCOVA) on average fixation duration showed that this difference was not significant. No between-subject effects of gesture counterbalance order were found, $F(1, 12) = 0.831$, $p = .380$, TOH type, $F(1, 12) = .09$, $p = .776$, and its interaction, $F(1, 12) = 0.66$, $p = .433$. Furthermore, average fixation duration was not significantly affected by gesture (estimated means in ms $= 300.25$, $SE = 15.16$, 95 % CI 267.20– 333.25) versus no gesture (estimated means in ms $= 284.80$, $SE = 14.96$, 95 % CI 252.23 – 317.42), $F(1, 12) = 0.25$, $p = .625$, nor was there an interaction effect of gesture and VPT, $F(1, 12) = 0.14$, $p = .716$.

Discussion

Prior research has shown that gesturing may compensate for high working memory load (e.g. Marstaller and Burianová 2013; Pouw et al. under review). However, it is not yet clear *how* gestures perform this cognitive function. The present study investigated the hypothesis that pointing gestures, by exploiting space, reduce the need for exploiting the visual presentation of the task in the external environment as a way to anchor mental simulations. Consequently, we expected less eye movements to be made when participants gestured during mental problem solving of the Tower of Hanoi (TOH) than when they did not gesture, because gestures can come to "stand in" for eye movements as an anchor for mental simulations in the external environment. That is, through pointing, gesturers can spatially index mental simulations of moving the discs from one peg to another in peri-personal space, rather than moving the eyes to project imagined disc movements onto the visual presentation of the task. Given that gestures can compensate for high cognitive load, we expected this effect to be stronger for those individuals who have lower visual working memory capacity (as problem solving places higher demands on their resources).

In line with this hypothesis, our results showed that gesturing lowered saccade counts during mental problem solving, and more strongly so for those with a lower visual working memory capacity. As such, this study makes a novel contribution towards explaining (one of) the mechanism(s) through which gestures may support (mental) problem solving. Whereas eye movements allow for projecting mental simulations in the external environment, gestures do this in exploiting peri-personal space through proprioceptive monitoring and peripheral visual control, thereby offloading visual working memory processes.

An important question is whether we can exclude that the effect of gesture on eye movements is an epiphenomenon, i.e. functionally irrelevant for mental problem solving? We think that gesture's effect on eye movements are not likely to be epiphenomenal as there are a host of findings which show that eye-movement patterns are crucial for thinking through the solution space of a problem (Spivey and Dale 2011) and to visual imagery in general (e.g. Brandt and Stark 1997; Johansson et al. 2006; Laeng and Teodorescu 2002). However, we do not (and cannot) claim (based on the present data) that reduction in saccade count is necessarily beneficial for problem solving as opposed to a more visually dominant strategy. However, given that eye movements are highly likely to be functionally relevant to mental simulations, and given the present findings that especially those gesturers with a lower visual working memory capacity considerably alter their

gaze patterns without significant loss in performance, it is likely that there is some trade-off mechanism present.

But what is the exact nature of this trade-off mechanism? Although this question cannot be definitively answered based on our data, the present study does suggest that the change from a visually dominant strategy to a strategy that exploits sensory consequences of gesture (especially proprioception) may offer a preliminary explanation. Recall that a visually dominant strategy involves moving the eyes in a way that corresponds with mentally moving the discs from one peg to another. This allows a way to anchor mental transformation on a visual presentation of the task (see Fig. 1, last frame). This strategy thus involves mental projection onto the external environment, where the external environment offers an anchor or reference that is meaningful to the task (e.g. Kirsh 2009). Pointing gestures can, we think, fulfil the same function as eye movements. However, pointing fulfils this function with different and less visually dominant resources. Namely, through pointing, peri-personal space is sequentially filled by positions of the hand that are, by physical human nature, monitored through proprioception and/or (peripheral) visual control (e.g. Bremner and Cowie 2013). The locations that the hand takes in space during pointing corresponds with the mental transformation being made by the gesturer. That is, mentally simulating the move of a disc corresponds to pointing from one location to another. The reason why we think pointing is not a visually dominant strategy, is that if participants' pointing gestures were actively visually tracking their pointing movements, then we would not have observed a difference in saccades between gesture versus no condition, as mental transformation in both cases are visually tracked (albeit in the gesture condition via an external loop). This was not the case. Furthermore, informal inspection of the videos reveals that participants were indeed not looking directly at their hands during gesturing. This leads us to our interpretation that gestures must provide some additional resource for spatially indexing mental transformations. We thus think that next to peripheral vision, proprioception can offer a natural way to monitor the hand in space as to spatially index mental transformations. Finally, although we cannot definitively establish that gestures are indeed proprioceptively dominant in this case, it does serve as an additional explanation of why those with a lower visual working memory capacity (a proxy for visual mental imagery ability) are especially likely to reduce their eye movements. Namely, those problem solvers that are prone to have difficulty projecting/simulating visual transformations on the environment, can reap the benefits of spatially indexing mental transformation in a non-visually dominant way through pointing (using the proprioceptive sense of the

hand in space). Findings that gesturing is especially potent for those with a lower working memory capacity (Marstaller and Burianová 2013; Pouw et al. under review), and is beneficial even when participants cannot see their own hands (Cooperrider et al. 2015), concur with this idea that switching to a non-visually dominant strategy is possible and perhaps potent for some but not all problem solvers.

Another question that could be raised is whether present results exclude a strict motor-based interpretation of gesture, wherein gestures effect should be attributed to reuse (internal simulations) of motor experience (Hostetter and Alibali 2008). Namely, a strict motor-based interpretation may entail that the *motor intention to produce a pointing gesture*, rather than the *actual bodily gesture and its sensory consequences*, activates/supports internal motor simulations which in some way affects gaze behavior as observed in the present study. This is in contrast to the embedded/extended approach which assumes that any explanation of a cognitive function of gesture must always lay (at least in part) in the sensory consequences of gesturing that are used in some cognitively potent way (Pouw et al. 2014). However, the present study was not designed to differentiate between these interpretations. Future research could focus on distinguishing a strict motor-based interpretation from an interpretation that emphasizes sensory consequences of gesture. This can be done by manipulating gesture intention (as to trigger motor simulations) versus actual gesture production. If the production of gestures plays no functional role in the present effect, then the intention to gesture should produce the same effect on eye movements (without loss in problem-solving performance). Finally note that the embedded/extended and motor-based approach can also be complementary. Under such a hybrid view, gestures arise out of motor simulations and have sensory consequences which further affect ongoing simulation-based cognitive processes (e.g. Pouw and Hostetter 2016).

Our study has limitations. First, it should be stressed that the current study is small in scale, and as such, definitive conclusions on the precise role of pointing on problem-solving processes should not be drawn from the present data. In particular, the present lack of an effect of gesture on problem-solving performance should be treated with caution as similar studies that did find a beneficial effect investigated this with a larger sample (e.g. Chu and Kita 2011; Garber and Goldin-Meadow 2002; Pouw et al. under review). That is, in contrast to our expectation, we did not find beneficial effects of gesturing during the mental problem-solving phase on TOH performance (TOH solving speed and solving steps). This is in contrast to prior findings (Pouw et al. under review), but important differences between the current and prior study lie in the design. First, participants in the prior study had more mental solving time before they physically performed the task: 150 versus

60 s in the present study. Second, whereas gesturing was a between-subject factor in the prior study, it was a within-subject factor in the present study. As such, even though it is unlikely given that the rule was inversed between tasks and the analysis of order effects revealed no significant differences, we cannot rule out entirely that there were carry-over effects that may have eliminated potential beneficial effects of gesturing on performance (especially since the number of participants per group in the order analyses were based was very small). For example, Chu and Kita (2011) have found that the beneficial effects of gesture can carry over to a subsequent task (similar in nature) when gesturing is prohibited. Additionally, it could be the case that pointing gestures are less beneficial for problem-solving performance as compared to co-speech iconic gestures that have been found to co-occur with *verbal* explanations of solving the TOH (e.g. Cook and Tanenhaus 2009; Cooperrider et al. 2015), wherein participants gesture as if grasping actual discs. Future research should further investigate whether iconic gestures during actual problem solving may have different effects than pointing gestures. For example, this can be done by letting participants verbally explain the solution of the TOH (e.g. Cooperrider et al. 2015). Yet there are several reasons why in the present case iconic gestures might not be particularly effective. Firstly, in a previous study (Pouw et al. under review), we have found that pointing gestures, but not iconic gestures, are spontaneously produced during mentally solving a physical Tower of Hanoi task without speech. This suggests that iconic gestures may be co-dependent on speech production and not naturally employed during mental problem solving without the additional constraint to verbalize one's thoughts. Furthermore, in this previous study, pointing gestures were found to benefit performance on subsequent solving of the TOH when cognitive load is high. Finally, the reason why iconic gestures are held to affect mental problem solving of the TOH is that they offer a correspondence with the actions to be performed on the actual task (Cook and Tanenhaus 2009; Cooperrider et al. 2015). Yet, in the present case, manipulation of the task was mouse-based, which does not correspond with a grasping action. In sum, although iconic gestures may offer unique or better cognitive support for problem solving, in the present non-verbal mouse-based task we doubt whether iconic gestures are more potent than pointing gestures.

A second limitation of the present study is that it relied on eye-movement frequency counts, and therefore does not yet illuminate the precise dynamics of pointing and gaze behaviour (i.e. when and how participants use gestures during [mental] problem solving and how this affects their eye movements). The benefit of our mobile eye-tracking device was that it allowed for maintaining natural degrees

of freedom in hand movement, which is more difficult to obtain (at present) with remote eye-tracking devices. Nevertheless, the higher temporal and spatial resolution that can be obtained with remote eye-tracking devices would allow us to address in more detail how eye movements are affected by gestures in future research.

Despite these limitations, this study made a first step towards explaining (one of) the mechanism(s) through which gestures may support (mental) problem solving. Our findings suggest that gesturing may provide a unique embodied resource, exploiting peri-personal space, which may come to stand-in for visually dominant strategies when these prove to be insufficient for meeting the cognitive demands imposed by the task. Taking gaze behaviour into account in future research may enhance our understanding of the role that non-communicative pointing gestures play in problem-solving processes, for individuals differing in cognitive dispositions.

References

Brandt SA, Stark LW (1997) Spontaneous eye movements during visual imagery reflect the content of the visual scene. J Cogn Neurosci 9(1):27–38. doi:10.1162/jocn.1997.9.1.27

Bremner AJ, Cowie D (2013) Developmental origins of the hand in the mind, and the role of the hand in the development of the mind. In: Radman Z (ed) The hand, an organ of the mind: What the manual tells the mental. MIT Press, Cambridge, pp 255–268

Cappuccio ML, Chu M, Kita S (2013) Pointing as an instrumental gesture: Gaze representation through indication. Humana. Mente: Journal of Philosophical Studies 24:125–149

Chu M, Kita S (2011) The nature of gestures' beneficial role in spatial problem solving. J Exp Psychol Gen 140(1):102–115. doi:10.1037/a0021790

Chu M, Meyer A, Foulkes L, Kita S (2013) Individual differences in frequency and salience of speech-accompanying gestures: the role of cognitive abilities and empathy. J Exp Psychol Gen 143:694–709. doi:10.1037/a0033861

Clark A (2013) Gesture as thought. In: Radman Z (ed) The hand, an organ of the mind: What the manual tells the mental. MIT Press, Cambridge, pp 255–268

Cook SW, Tanenhaus MK (2009) Embodied communication: speakers' gestures affect listeners' actions. Cognition 113(1):98–104. doi:10.1016/j.cognition.2009.06.006

Cooperrider, K., Wakefield, E., & Goldin-Meadow, S (2015). More than Meets the Eye: Gesture Changes Thought, even without Visual Feedback. *Proceedings of the 37th Annual Meeting of the Cognitive Science Society*. Austin, TX: Cognitive Science Society

Delgado B, Gómez JC, Sarriá E (2011) Pointing gestures as a cognitive tool in young children: experimental evidence. J Exp Child Psychol 110:299–312. doi:10.1016/j.jecp.2011.04.010

Della Sala S, Gray C, Baddeley A, Wilson L (1997) The visual patterns test: A new test of short-term visual recall. Bury St. Edmunds, Thames Valley Test

Garber P, Goldin-Meadow S (2002) Gesture offers insight into problem-solving in adults and children. Cognitive Science 26(6):817–831. doi:10.1016/S0364-0213(02)00087-3

Johansson R, Holsanova J, Holmqvist K (2006) Pictures and spoken descriptions elicit similar eye movements during mental imagery, both in light and in complete darkness. Cognitive Science 30(6):1053–1079. doi:10.1207/s15516709cog0000_86

Kirsh D (2009) Projection, problem space and anchoring. Proceedings of the 31st annual conference of the cognitive science society. Lawrence Erlbaum Associates, Hillsdale, pp 2310–2315

Laeng B, Teodorescu DS (2002) Eye scanpaths during visual imagery reenact those of perception of the same visual scene. Cognitive Science 26(2):207–231. doi:10.1016/S0364-0213(01)00065-9

Lee IA, Preacher KJ (2013) Calculation for the test of the difference between two dependent correlations with one variable in common [Computer software]. Available from http://quantpsy.org. Accessed 17 Mar 2016

Logan T, Lowrie T, Diezmann CM (2014) Co-thought gestures: supporting students to successfully navigate map tasks. Educational Studies in Mathematics 87:87–102. doi:10.1007/s10649-014-9546-2

Marstaller L, Burianová H (2013) Individual differences in the gesture effect on working memory. Psychon Bull Rev 20:496–500. doi:10.3758/s13423-012-0365-0

Patsenko EG, Altmann EM (2010) How planful is routine behavior? A selective-attention model of performance in the Tower of Hanoi. J Exp Psychol Gen 139(1):95. doi:10.1037/a0018268

Pouw WTJL, Hostetter AB (2016) Gesture as predictive action **(under review)**

Pouw WTJL, De Nooijer JA, Van Gog T, Zwaan RA, Paas F (2014) Toward a more embedded/extended perspective on the cognitive function of gestures. Frontiers in Psychology 5:359. doi:10.3389/fpsyg.2014.00359

Pouw, W. T. J. L., Eielts, C., van Gog, T., Zwaan, R. A., & Paas, F. (under review). When gestures are put into thought: The role of co-thought gestures, visual working memory, and task complexity in problem-solving

Smithson L, Nicoladis E (2014) Lending a hand to imagery? The impact of visuospatial working memory interference upon iconic gesture production in a narrative task. J Nonverbal Behav 38(2):247–258

Spivey MJ, Dale R (2011) Eye movements both reveal and influence problem solving. In: Liversedge SP, Gilchrist ID, Everling S (eds) The Oxford handbook on eye movements. Oxford University Press, Oxford, pp 551–562

Thomas LE, Lleras A (2007) Moving eyes and moving thought: on the spatial compatibility between eye movements and cognition. Psychon Bull Rev 14(4):663–668. doi:10.3758/BF03196818