Original Papers

Polish Psychological Bulletin
2019, vol. 50(2) 184–192
DOI - 10.24425/ppb.2019.129451

*Agata Złotogórska-Suwińska**
*Adam Putko**

Visual Perspective-Taking: Are Theory of Mind and Executive Function Involved?

***Abstract**: The aim of the study was to examine the relationships of Level-1/Level-2 visual perspective-taking (VPT-1/VPT-2) with theory of mind (ToM) and executive function (EF). Seventy-six adults aged 18 to 48 years participated in the study. To compare the relationships of the two levels of perspective-taking with the aforementioned abilities, the same stimuli were used in both Level-1 and Level-2 trials of the VPT task. ToM abilities were evaluated with the Strange Story task, and EF using the TMT and WCST tests. It was found that controlling for age-related differences, VPT-1 was not associated with either ToM or such components of EF as executive control and set-shifting. VPT-2 was positively related to ToM, but it was not related to EF. The relationship between VPT-2 and ToM was specific, not mediated by domain-general processing capabilities. The obtained results provide further evidence to support the view that distinct mechanisms underlie Level-1 and Level-2 perspective-taking.*

***Keywords**: visual perspective-taking, theory of mind, executive function*

Introduction

Taking visual perspective of another person is one of the important abilities used in our daily interactions with other people. It involves assigning perceptual mental states to others (e.g., ‘she sees it’ or ‘she sees it differently than I do’), which may become the basis for reasoning about other kinds of mental states (e.g., knowledge, beliefs, desires, etc.) of people whose behaviour we want to understand or predict. The majority of previous studies on visual perspective-taking (VPT) focused on determining the age at which this ability develops in children, changes in its level associated with cognitive ageing or disorders of its development. Less attention has so far been devoted to the problem of relations that occur among VPT and other social and cognitive abilities. First of all, it is not clear whether VPT is associated with theory of mind (ToM) – the ability to ascribe mental states to others and to oneself in order to predict or explain someone’s behaviour. According to some researchers (e.g., Aichhorn, Perner, Kronbichler, Staffen, & Ladurner, 2006), such a relationship occurs because at the root of both VPT and attributing mental states is a certain common factor – an understanding of perspective. Second, because ToM abilities are usually associated with the level of executive function (EF) – processes involved in the conscious control of thought and action (e.g., Zelazo & Müller, 2002) – the question arises whether a similar relationship also occurs in the case of VPT. In the following section of the article, the current state of the research regarding the aforementioned issues will be outlined. Next, the results of our own study will be presented, the purpose of which was to examine the relationship of VPT with ToM abilities and EF.

Visual perspective-taking and theory of mind

The interest in VPT abilities derives from the work of Piaget and Inhelder (1948/1956) on children’s perception of space. Using the task called the Three Mountains problem, Piaget has found that children under the age of 7 who are still in the preoperational stage of thinking usually cannot solve this task. When they were to pick out from the set of pictures the one which is most suited to the view of the mountains seen by the doll occupying a position in space other than them, they chose most often the picture depicting the mountains from their own vantage point. They were unable to conceptualize what the scene looks like from someone else’s perspective,

* **Institute of Psychology Adam Mickiewicz University in Poznań**

Corresponding author: Agata Złotogórska-Suwińska, e-mail: zlotogorska@gmail.com

because according to Piaget, they were centred on their own point of view (Wood, 2006). And although later research showed that Piaget did not appreciate children's abilities in this area, his research became a source of inspiration for other researchers to introduce an important distinction between the two levels of perspective-taking: Level-1 and Level-2 (Flavell, Everett, Croft, & Flavell, 1981). Level-1 perspective-taking is based on the understanding that objects visible to one person are not necessarily visible to the other. On the other hand, Level-2 perspective-taking comprises an understanding that the same object or scene can be viewed differently by a person occupying a place other than us in space. Current research shows that the ability to take Level-1 perspective (VPT-1) develops around the age of 2 years (e.g., Moll & Tomasello, 2006), while the ability to take Level-2 perspective (VPT-2) develops around age 4, that is, at the age when children also begin to pass an explicit test of false belief understanding.

A neuroimaging study by Aichhorn and colleagues (2006) provides a strong basis for predicting that VPT abilities are associated with ToM. It showed that during performance of the VPT-2 task, the same area of the brain is active – the temporo-parietal junction (TPJ) – which is also involved, as research shows (e.g., Saxe & Kanwisher, 2003), in ToM tasks, especially those that require the attribution of false beliefs. According to Aichhorn and colleagues, the dorsal part of the TPJ region is responsible for representing perspective differences – the ability underlying both Level-2 perspective-taking and false beliefs ascription. They argued that only Level-2 perspective-taking, but not Level-1, requires the understanding that the same object or scene may be perceived differently by a person occupying a place other than us in space. On this basis, it can be presumed that of the two levels of VPT, Level-1 and -2, Level-2 will be associated with ToM abilities that require a similar 'perspectival' understanding, i.e., an understanding that one and the same state of affairs can be differently conceived by two persons (cf. Perner, Brandl, & Garnham, 2003). However, VPT-1 as being deprived of such requirements should not be related to ToM abilities, or at least this relationship should be weaker than in the case of VPT-2.

The results of previous studies, however, do not provide an unambiguous answer to the question of whether there is a relationship between VPT and ToM abilities. In Langdon and Coltheart's study (2001, exp. 2), conducted on a group of young adults, a correlation ($r = -.49$) was found between response times on the VPT-2 task and ToM abilities (measured by a picture story task that required reasoning about false beliefs), but only in such a condition of this task which required participants to imagine moving themselves to another viewer position from which they were to make judgements about the scene presented to them. However, in the condition that required participants to imagine rotating not their body, but the visual scene, no relationship was found between performance on this task and ToM abilities. The relationship between VPT-2 and ToM was also found, among others, in Białek's study (2011), in which Brüne's Picture Sequencing Task was used as a measure of ToM.

Important data are also provided by research on atypical development. As the review of studies on VPT abilities in ASD indicates (Pearson, Ropar, & de Hamilton, 2013), individuals with ASD usually have no difficulty with intentional, explicit Level-1 perspective-taking, while it is difficult for them both to adopt a Level-2 perspective and to infer mental states. However, there are also such studies whose results point to the lack of a relationship between VPT-2 and ToM abilities. In a study by Peterson, Peterson and Webb (2000) with blind children between the ages of 6 and 12, the majority had no difficulty in the tactile versions of VPT-2 tasks, but most of those who were 9 years of age or younger did not succeed at the false belief tasks. A similar pattern of results was found by Peterson (2003) in deaf children aged 4–13 years, growing up in hearing families. The vast majority of these children passed the tests of VPT-2, while only about a quarter of them passed the tests involving false belief tasks. In conclusion: previous research does not provide a definite answer to the question of whether the ability to take a visual perspective of another person is associated with ToM abilities. The limitation of this research is also the lack of applying appropriate conditions or control tasks that would check whether the possible relationship between performance on VPT and ToM tasks is specific and not mediated by certain other requirements that both of these tasks share, e.g., related to their visual or spatial nature.

Visual perspective-taking and executive function

Taking someone else's visual perspective requires suppressing one's own perceptual experience, refraining from thinking about what one sees and switching to the perspective of another person. In this view, perspective-taking is a process which is likely to involve inhibitory control and set-shifting – key components of EF. The role of EF in processing a visual perspective of another person has so far been the subject of only a few studies. In one of the most important works of research on this issue, Qureshi, Apperly and Samson (2010) used the dot-perspective task designed by Samson and colleagues (2010) to measure VPT-1, in which the participants were asked to make speedy judgments about how many dots they could see on the walls of the presented room (self-perspective trials) or how many dots another person standing in this room could see (other-perspective trials). The task, therefore, required participants to switch in a long sequence of trials from their own perspective to someone else's and vice versa. To assess the extent to which EFs are involved in VPT-1, Qureshi and colleagues employed a dual-task paradigm. In the dual-task condition, participants performed the dot-perspective task concurrently with Luria's tapping task, involving inhibition. The authors assumed that if VPT-1 utilizes automatic processes not involving EF, then simultaneously performing Luria's task should not affect performance on the VPT-1 task. The study showed, however, that in the dual-task condition, compared to the single-task condition, participants generally incurred higher processing costs, responding more slowly and making more errors, in the judgments they made both about what the other person saw and about what they themselves saw. Qureshi and colleagues interpreted this finding as evidence of the involvement of EF (mainly inhibition), not so much in the process of 'calculating' someone else's perspective, but in selecting an appropriate perspective (attenuating one perspective and switching to another).

The involvement of EF in perspective-taking is also suggested by research on cognitive ageing. Studies of people above the age of 50 indicate that with increasing age the ability to adopt the Level-2 perspective deteriorates, which is reflected in the increase in the number of egocentric errors (e.g., McDonald & Stuart-Hamilton, 2003). It is assumed that the decline of this ability in people at this age may be to a certain extent caused by a decrease in EF – cognitive flexibility and inhibition.

The limitation of previous studies on the role of cognitive and social processes in Level-1/Level-2 VPT is that they use tasks that are not always fully comparable, differing from each other in some respects, such as the type of test stimuli and the way of presenting or responding to them, which in themselves may impose different performance requirements, making it difficult to compare the results and draw stronger conclusions. Therefore, in the current study concerning both levels of perspective-taking, we decided in the VPT task to apply trials of an almost identical nature, which used the same test stimuli, but of course differed in the test questions. This allowed for a closer comparison of relations between the two levels of perspective-taking and ToM abilities, as well as selected components of EF.

The current study

The aim of our study was to answer the question of whether both levels of perspective-taking are related to ToM abilities, as well as to EF. Based upon the literature review, we hypothesized that VPT-2, but not VPT-1, is related to ToM abilities. We were also interested in whether both levels of perspective-taking differ from each other in terms of relationships with EF, but due to the exploratory nature of this part of our research, we have not put forward specific hypotheses.

In our study, we focused on the assessment of cognitive flexibility to check whether this kind of EF is related to the efficiency of perspective-taking. It can be assumed that switching between "Self" and "Other" perspectives requires suppressing one perspective in order to adopt the other. The blocks of trials in our perspective-taking tasks were of mixed nature – they consisted of both "Self" and "Other" trials presented in an alternating manner. So these trials required participants to shift their task-sets to flexibly switch from one perspective to another. We were particularly interested in whether both VPT tasks, Level-1 and Level-2, pose similar requirements related to cognitive flexibility. To assess individual differences in this kind of EF we decided to use the Trail Making Test (TMT). The main reason for using this test was that, as shown by Arbuthnott and Frank (2000), performance on part B of the TMT is associated with set-switching costs, and the derived B:A ratio score in this task can provide relatively clear indication of executive control needed to resolve suppression of a previously-abandoned task. The additional reason why we chose TMT is that we were also interested in assessing the extent to which performance on both VPT tasks depends on domain-general processes. Given that the main measure of performance on VPT tasks are response times, we decided to check the relationship between them and the processing speed, which can be treated, along with working memory capacity, as one of the most important central cognitive resources (e.g., Salthouse, 1996). So, we used part A of the TMT to evaluate individual differences in the processing speed.

The task employed to assess VPT abilities was modelled on the procedure developed by Samson et al. (2010). In order to evaluate the participants' ability to adopt not only the Level-1 but also the Level-2 perspective, unlike Samson and colleagues, we applied three-dimensional objects known to everyone from everyday life as the stimulus material. By using the same stimuli, although of course different test questions, in the task to measure both levels of perspective-taking, it was possible to directly compare the relationships between performance on this task and other abilities we were interested in – ToM and EF.

Method

Participants

Seventy-six volunteers (including 41 women) aged 18 to 48 participated in the study ($M = 29.14$, $SD = 8.32$). Almost half of the participants (47%) had higher education, 15% post-secondary education, 30% secondary, 5% vocational and 3% basic. All participants had normal or corrected-to-normal vision and gave informed consent to participate in the study.

Materials

Visual perspective-taking task

It was a computer task prepared in the DMDX program (Forster & Forster, 2003) by that study's lead author. Response times and accuracy of responses to stimuli were measured. The stimuli were photos of a person (a man or a woman) sitting at a table and looking at the object on it. Four objects were used that looked different from every side: headphones, a kettle, a teddy bear, and an iron. There were four kinds of trials, differing in the level of perspective (*Level-1/Level-2*) and the type of perspective (*Self/Other*) that participants should adopt when assessing the visual scene. Level-1 trials (14 trials) proceeded as follows. First, a test question was displayed on the computer screen indicating whose perspective the participant should take: another's person ('Does he/she see the object?') or one's own ('Do you see this object?'). Below the test question, there was a photo of the object (a cue) about which the participant was asked, presenting this object from one of its four sides. After the participant pressed the space bar, a fixation point appeared for

Figure 1. Example of Level-1 Other trial in the VPT task

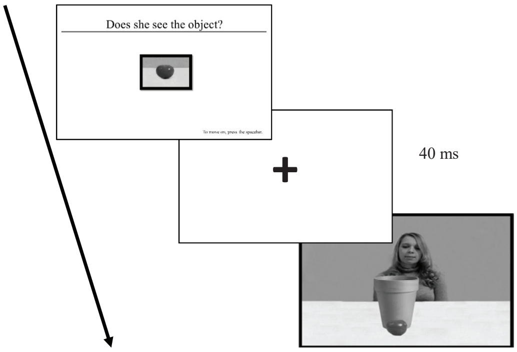
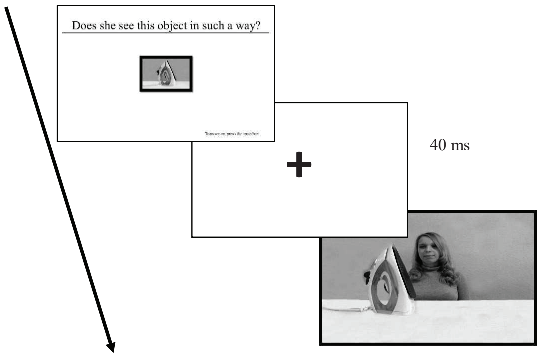


Figure 2. Example of Level-2 Other trial in the VPT task



400 ms on a blank screen, placed at a position where, in the photograph that was presented a moment later, a person's eyes were shown. After this time, for three seconds, a photograph of a person looking at one of the two objects on the table was presented. The participant's task was to answer the test question by pressing the green 'YES' key or the red 'NO' key as soon as possible. Level-2 trials (16 trials) proceeded in the same manner as those of Level-1, but they differed in the test question (see Figure 1 and 2). On Other-perspective trials, the participant judged whether the person presented in the photo sees the object in the same way as the photo-cue showed it ('Does he/she see this object in such a way?'). On Self-perspective trials, the participant judged whether he or she saw the object currently presented in the picture in the same way as the photo-cue showed it ('Do you see the object in this way?'). In addition to the test trials, we used 8 'filling' trials with an object that looked the same on every side (a lass, a tomato, a flowerpot, or a dish), in which the test question concerned the Level-1 perspective. The inclusion of these trials was aimed at preventing the participants from developing the belief that the objects shown would always look different on each side. Responses in filling trials were recorded but not analysed, due to the much simpler nature of these types of stimuli compared to those shown in the test trials. The order of the presentation of all trials was random, with the one limitation that no more than two trials of the same type could occur in a row. In half of all the trials presented in the task, the correct answer was 'YES', and in half of them – 'NO'. The reliability of the measurement in the VPT task was good or high (Level-1 Self trials: Cronbach's $\alpha = .90$; Level-1 Other $\alpha = .94$; Level-2 Self: $\alpha = .84$; Level-2 Other: $\alpha = .91$).

Strange Stories Task (Happé, 1994)

Eight stories from the Strange Stories collection were used. These were four stories that required participants to justify someone's behaviour in terms of mental states (mentalistic stories: two first-order ToM stories [a lie, a white lie]; and two second-order ToM stories [lost glove, ice cream bus]) and four control stories that did not require referring to mental states to explain someone's behaviour (a burglar in a jewellery store, payment in instalments, meringues, a lost hat). Answers in six stories were scored from 0 to 2 points (0 = incorrect answer or no response, 1 = correct but incomplete answer, 2 = complete answer), two others (one mentalistic, two control) from 0 to 1. The interrater agreement ranged from .44 to .98.

Trial Making Test (TMT; Reitan, 1971)

TMT consists of two parts: A and B. In part A, the participant's task is to connect 25 encircled numbers randomly distributed on a sheet of paper in ascending order by drawing a pencil line. The test part of the task is preceded by a practice part, in which the numbers from 1 to 8 should be connected. In part B the task is the same, except that encircled numbers (from 1 to 13) and letters (from A to L) should be connected in alternating number-letter order, i.e., 1-A-2-B-3-C etc. In a practice part, there are circles with numbers from 1 to 4 and letters from A to D. If the participant makes an error, the participant is instructed to return to the place where the error was made and continue. The score on each part of the TMT represents the amount of time required to complete the task. Completion time on part A is considered as a measure of processing speed, while completion time on part B is seen as an indicator of the ability to shift between cognitive sets (cognitive flexibility) and working memory (Sánchez-Cubillo et al., 2009). In the analysis of the results, the B : A ratio was also calculated as a measure of executive control (e.g., Arbuthnott & Frank, 2000). The lower the value of this indicator, the higher the level of executive control.

Wisconsin Card Sorting Test (WCST; Berg, 1948)

A computerized version of the WCST was employed, prepared in Inquisit 4 Lab version 4.0.5.0. According to the standard procedure, the participants had to sort the 128 response cards, matching them to one of the four stimulus cards according to a rule that they had to discover through feedback ('correct' or 'incorrect') provided by the computer. The sorting rules were the following: colour, shape, and number, each of which occurred twice. After a participant correctly sorted 10 consecutive cards, the sorting rule changed without informing the participant. The task was continued until either all 128 cards were sorted or six categories were successfully completed. The number of perseverative responses was used in the present study as a measure of cognitive flexibility.

Procedure

All participants were tested individually in one session. Due to the correlational study design, the tasks were presented in a fixed order. The first task administered was the TMT; the next was the VPT task displayed on a 15-inch computer screen. The test part of the VPT task was preceded by a series of 14 practice trials, in which the participant had to respond correctly on at least 10 trials so that he or she could continue the task. All participants successfully passed a practice part. The last two tasks were the Wisconsin Card Sorting Test and the Strange Stories task. At the end, participants completed a demographic data survey.

Results

Preliminary analyses showed no significant differences in the results obtained by gender. When analysing response times, incorrectly answered trials were removed. In the case of several variables, moderate or weak correlations with age were found. This was so in the case of ToM stories, $r(76) = -.26$; response times on Level-1 Other trials, $r(76) = .34$; Level-2 Other trials, $r(76) = .36$; and two variables from the TMT: part A, $r(76) = .26$; and the B : A ratio, $r(76) = -.27$. Variables whose distributions differed from normal (i.e., those for which the ratio of skewness to standard error was lower than –2 or higher than 2) were subjected to appropriate transformations to improve their distributions. The descriptive statistics of the analysed variables presented in Table 1 refer to their values before they had been possibly transformed.

As shown in Table 1, the average percentage of correct responses in the VPT task, both Level-1 and Level-2, was very high and ranged from nearly 80 to over 94%. In the VPT-1 task, the percentage of correct responses on Self trials did not differ significantly from that on Other trials, $t(75) = 1.43$, $p > .05$, Cohen's $d = 0.33$. However, the average response time on Self trials was significantly lower than on Other trials, $t(75) = 5.83$, $p < .001$, Cohen's $d = 1.35$. In the VPT-2 task, the average percentage of correct responses on Self trials was significantly higher than on Other trials, $t(75) = 3.35$, $p = .001$, Cohen's $d = 0.77$. The average response time on Self trials of this task was significantly lower than on Other trials, $t(75) = 6.64$, $p < .001$, Cohen's $d = 1.53$. As regards performance on TMT, the average completion time of part A was significantly lower than part B, $t(75) = 7.04$, $p < .001$, Cohen's $d = 1.63$. It should be noted that the high average percentage of correct responses on Self and Other trials of the VPT-1 task was very similar to that in Samson et al. (2010), where it exceeded 90%. Regarding the VPT-2 task, the average percentage of correct responses was only slightly lower in the present study than in the study by Surtees, Butterfill and Apperly (2012), who used a similar procedure and found that task performance in the group of adults exceeded 90% on both kinds of trials. In further analyses, as an indicator of performance on the VPT task, only response times were used, due to the high percent of correct responses on this task.

Table 1. Descriptive Statistics for Scores on the Strange Stories, TMT, WCST, and Visual Perspective-Taking Task

Task	*M*	*SD*	Range
Strange Stories			
ToM stories	4.63	1.62	1–7
Control stories	4.18	1.68	0–7
TMT			
A (sec)	64.86	17.17	31–121
B (sec)	80.51	24.29	41–191
B:A	1.27	.30	.75–2.00
WCST			
Perseverative responses (%)	11.66	5.25	0–30.49
Level-1 VPT			
Self %	92.11	12.01	67.00–100
Other %	94.41	8.01	75.00–100
Self RT (ms)	651	195	329–1111
Other RT (ms)	742	188	417–1096
Level-2 VPT			
Self %	87.99	12.98	50–100
Other %	79.93	16.97	38–100
Self RT (ms)	719	203	350–1,200
Other RT (ms)	866	251	390–1,440

Note. TMT = Trail Making Task; ToM = theory of mind; VPT = visual perspective task; WCST = Wisconsin Card Sorting Test.

Level-1 visual perspective-taking

As shown in Table 2, of the two measures of VPT-1, only the response times on Other trials correlated negatively with ToM scores and positively with TMT-A. To check whether these variables are unique predictors of performance on the VPT-1 task, a regression analysis was conducted using the bootstrap method. In this analysis, the influence of age was also controlled for because it correlated positively with response times on Other trials of the VPT-1 task and with TMT-A, as well as negatively with performance on ToM stories. The significance of the regression coefficients was tested on the basis of the 10,000 bootstrapped samples. A regression analysis showed that only age significantly predicted response times on Other trials of the VPT-1 task, scores on ToM stories were a marginally significant ($p < .1$) predictor, and TMT-A was not a significant predictor (see Table 3). These findings lead to the conclusion that VPT-1 is not related either to

Table 2. Correlations (Pearson's r) Among the Measures of Visual Perspective-Taking, Theory of Mind, and Executive Function

	1.	2.	3.	4.	5.	6.	7.	8.	9.
1. L-1 Self									
2. L-1 Other	.75**								
3. L-2 Self	.66**	.78**							
4. L-2 Other	.48**	.76**	.66**						
5. ToM stories	–.03	–.29*	–.22#	–.36**					
6. Ctrl stories	–.19	–.03	–.10	.04	.18				
7. TMT-A	.19	.25*	.28*	.35**	–.27*	–.10			
8. TMT-B	.14	.17	.08	.09	–.16	–.04	.53**		
9. TMT B:A	–.08	–.15	–.21#	–.23*	.16	–.04	–.38**	.41**	
10. WCST	–.03	.01	.03	.05	–.14	–.04	.06	.13	.06

Note. Ctrl = control; L-1/L-2 = Level-1/Level-2; TMT = Trail Making Task; ToM = theory of mind; WCST = Wisconsin Card Sorting Test.

# $p < .1$; * $p < .05$; ** $p < .01$ (two-tailed).

Table 3. Regression Analysis (Final Model) for Response Times on Other Trials of the Level-1 VPT Task

Variable	*B*	*SE*	*p*	95% *CI*	sr^2
Age	5.67	2.56	.031	.57; 10.64	.057
ToM stories	–2.54	1.51	.090	–5.24; .66	.031
TMT-A	1.52	1.16	.180	–.62; 3.95	.018

Note. sr^2 = squared semi-partial correlation.

ToM abilities[1] or to the kind of EF that was measured in our study.

Level-2 visual perspective-taking

In the case of the VPT-2 task, more significant relationships were found (see Table 2). First of all, response times on Other trials of this task correlated negatively with scores on ToM stories, positively with TMT-A and the TMT B:A ratio. In contrast, response times on Self trials of these task correlated significantly only with TMT-A, whereas correlations of this variable with ToM stories and the TMT B:A ratio did not reach statistical significance ($p < .1$). To check whether the variables correlating with performance on Level-2 Other trials are independent predictors of it, a regression analysis was conducted. In this analysis, the influence of age was also controlled for, since it correlated positively with response times on Level-2 Other trials and TMT-A and negatively with ToM stories. First, we applied a stepwise regression method to select the best predictors. The analysis showed that three variables predicted response times on Level-2 Other trials: age, ToM stories, and TMT-A. However, the TMT B:A ratio did not contribute significantly ($p = .657$) to the increase in the variance explained of the dependent variable; thus, it was not included in the final model. The full model was well fitted to the data and accounted for 25% of the variance of the dependent variable, $F(3, 72) = 7.86$; $p < .001$; $R^2 = .25$; R^2 adjusted = .22. Next, we used the bootstrap regression analysis, the same as before, to estimate the significance of regression coefficients. It appeared that all three variables remained significant predictors of response times on Level-2 Other trials (see Table 4).

Table 4. Regression Analysis (Final Model) for Response Times on Other Trials of the Level-2 VPT Task

Variable	*B*	*SE*	*p*	95% *CI*	sr^2
Age	7.09	3.27	.043	.25; 13.91	.049
ToM stories	–4.36	1.99	.026	–7.93; –.36	.050
TMT-A	3.28	1.59	.045	.40; 6.79	.045

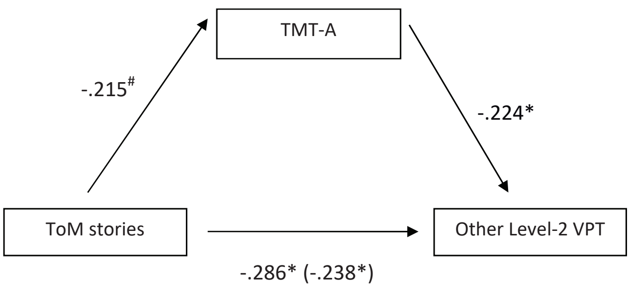
Note. sr^2 = squared semi-partial correlation.

[1] Following the suggestion of one of the anonymous reviewers, we also checked whether the pattern of correlations with the VPT and EF measures differs for the first- and second-order ToM stories. To this end, we calculated separate sums of scores for the first- and second-order ToM stories. It appeared that the second-order ToM correlated somewhat more strongly (–.31**) than the first-order ToM (–.28*) with Level-2 Other VPT. Of the remaining correlations, none achieved significance, and the strongest of them were marginally significant correlations of second-order ToM with Other Level-1 VPT (–.22) and TMT-A (–.20). Thus, the correlations for the second-order ToM seem to be stronger than for the first-order ToM, but probably a narrow range of such a partial variable prevents it from revealing strong enough relationships with other relevant variables.

Mediation analysis for the relationship between ToM scores and response times on Level-2 VPT task

It should be noted that the completion time of part A of the TMT task correlated negatively with scores on ToM stories and positively with response times on Level-2 Other trials (see Table 2). Such a pattern of correlations suggests that the relationship between Level-2 Other trials and ToM stories may be mediated by TMT-A. To check whether there is a mediation relationship between Level-2 Other trials and ToM stories, a mediation analysis was conducted using the PROCESS ver. 3.0 macro (Hayes, 2018). The mediation model was tested, in which scores on ToM stories were the independent variable, response times on Level-2 Other trials were the dependent variable, and TMT-A was the mediator variable. The model also included age as a covariate because, as mentioned previously, it correlated with response time on Level-2 Other trials and with other variables included in the model. As shown in Figure 3, when age was controlled for, the standardized regression coefficient for the relationship between ToM stories and TMT-A turned out to be only marginally significant ($p = .064$), while the standardized regression coefficient for the relationship between TMT-A and Level-2 Other trials remained significant ($p = .042$). The standardized indirect effect was (–.215) * (.224) = –.048. The significance of the indirect effect was tested using bootstrapping procedures. The unstandardized indirect effect calculated on the basis of the 10,000 bootstrapped samples was –.833, and the 95% confidence interval ranged from –2.759 to .056, indicating that the indirect effect was not significant. However, the direct effect between ToM stories and Level-2 Other trials, controlling for TMT-A and age, was still significant ($p = .032$). The results of the mediation analysis lead to the conclusion that the relationship between ToM stories and response times on Level-2 Other trials is not mediated by TMT-A.

Figure 3. Mediation model for the relationship between scores on ToM stories and response times on Other trials of the Level-2 VPT task



The figure shows the standardized regression coefficients for the relationship between ToM stories and Other trials as mediated by TMT-A, controlling for age. The standardized regression coefficient for the relationship between ToM stories and Other trials, controlling for age and TMT-A, is shown in parenthesis.

# $p < .1$; * $p < .05$.

Discussion

The first goal of the current study was to check whether there are relationships between Level-1/Level-2 visual perspective-taking and ToM abilities. We hypothesized that VPT-2, as opposed to VPT-1, is associated with ToM abilities. The study showed that performance on the VPT-2 task (response times on Other trials) correlated positively with ToM abilities (scores on ToM stories), even if age-related differences were taken into account. As regards VPT-1, it was also associated with performance on ToM stories; however, this relationship ceased to be significant when age was controlled for. The existence of a link between VPT-2 and ToM abilities is in line with the claim of Aichhorn and colleagues (2006) that at the root of both of these capabilities lies a common factor – an understanding of perspective. Understanding that the same state of affairs can be interpreted differently by different people, noticing differences in perspectives, is particularly important in ToM tasks that require the assignment of beliefs. Indeed, three out of the four mentalistic stories used in our study required the assignment of beliefs; in the fourth of them (white lie), one's belief was not the only or most important mental state that should be considered, but nor was this story devoid of requirements of this type. However, the lack of specific link between ToM abilities and VPT-1 is consistent with the assumption shared by many researchers that this type of perspective-taking does not require imagining what the other person sees, but consists in drawing a line of the person's sight and checking if the given object lies on this line.

The second important result obtained in the present study concerns Self trials in the VPT-1 task. In this type of trial, another person looking at the same scene from his or her own vantage point was in the field of view of the participant. Samson and colleagues (2010) found that in such trials participants spontaneously adopt the perspective of the other person. We were interested in whether performance on trials of this type is related to ToM abilities. It turned out that there is no such relationship. Three interpretations of this result are possible. The first, a spontaneous Level-1 perspective-taking is based on the same mechanism – following the line of sight – on which the explicit, intentional perspective-taking is based and does not require ToM abilities. The second possibility is that a spontaneous adoption of the other person's perspective involves ToM abilities, but it is a different type of abilities than those measured by the Strange Story task. In other words, the lack of a relationship results from the fact that performance in the first task is based on implicit ToM, and in the second on explicit ToM. This corresponds to the division of ToM abilities into two processing systems in Apperly and Butterfill's theory (2009): an efficient but inflexible implicit system and a cognitive demanding but flexible explicit system. A third possibility is that spontaneous Level-1 perspective-taking occurs only in the case of relatively simple stimuli, such as dots, as in the study by Samson and colleagues (2010), and does not occur in the case of more complex stimuli, such as those used in our study. The construction of our

VPT task does not allow for verifying the third explanation. It should also be noted that Level-2 Self trials were not associated with ToM abilities in our study. This result can be interpreted as being in line with Surtees and colleagues' study (2012), in which a spontaneous adoption of Level-2 perspective was not found in Self trials. According to the authors, this level of perspective-taking involves cognitively demanding processes and therefore does not occur spontaneously.

The second objective of the present study was to investigate the relationships between Level-1/Level-2 perspective-taking and EF. We were interested in whether both levels of this ability are related to such components of EF as set-shifting and executive control, but due to the exploratory nature of this part of the study, we did not form specific hypotheses. The study showed that measures of EF – the TMT B:A ratio and the percentage of perseverative responses in the WCST – do not correlate with response times on Level-1 trials. This result suggests that VPT-1 does not involve those components of EF that were measured in our study: set-shifting and executive control. This is consistent with the claim that Level-1 perspective-taking is an automatic process, as suggested by some studies (e.g., Michael et al., 2018; Qureshi et al., 2012), and is associated with the operation of a cognitively efficient but relatively inflexible system (Apperly & Butterfill, 2009).

Regarding Level-2 perspective-taking, we found that this ability correlated with a measure of executive control (the TMT B:A ratio), but the relationship was no longer significant when age and a measure of processing speed (TMT-A) were taken into account. There are two possible interpretations of this finding. According to the first interpretation, either VPT-2 does not involve executive control, or else the contribution of this component of EF to this kind of perspective-taking is relatively small, and hence difficult to detect in a correlational study based on a sample of such size as in the current study. The second interpretation refers to the conclusions drawn from a study by Samson, Houthuys and Humphreys (2015) carried out on people with lesions of various prefrontal cortex areas. This study showed that the ability to control interference related to its own perspective (egocentric tendency) is independent of the interference control revealed in the classical EF tests. According to the authors, this puts into question the domain-general nature of executive processes involved in perspective-taking. The results of our study, indicating the lack of a specific relationship between the ability to take a perspective of another person and EF, are in line with this conclusion.

An important result of our study is also that the relationship between ToM abilities and Level-2 perspective-taking (Other trials) is not mediated by general processing capabilities (processing speed measured by TMT-A), as demonstrated by mediation analysis. This finding fits well the ongoing discussion on the mechanisms underlying perspective-taking and ToM abilities. For example, it is argued that the alleged phenomenon of spontaneous adoption of Level-1 perspective is the result of non-specific, directional features of the stimulus, which in the dot-perspective task is a human figure facing one of the two sides (e.g., Santiesteban, Catmur, Hopkins, Bird, & Heyes, 2014). Our study indicates that although some of the processes involved in VPT-2 are probably domain-general (the relationship of the Level-2 Other trials with the processing speed[2]), another portion of them, as shown by mediation analysis, is specifically related to the attribution of mental states.

In future research, it would be worth checking if Level-2 perspective-taking is related only to the understanding of such mental states as beliefs, or also to other mental states, e.g., those more affective. Demonstrating that this type of perspective-taking is not associated with the understanding of mental states other than belief or belief-like states would provide further evidence for the claim that at the root of Level-2 perspective-taking and ToM is the ability to understand differences in cognitive perspectives. The results of the current study also suggest that the involvement of EF in Level-2 perspective-taking is probably low. However, it would be advisable to confirm this finding in an experimental study design, instead of a correlational one, and also applying other measures of EF.

Declaration of Conflicting Interests

The authors declare no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

References

Aichhorn, M., Perner, J., Kronbichler, M., Staffen, W., & Ladurner, G. (2006). Do visual perspective tasks need theory of mind? *Neuroimage*, *30*, 1059–1068. doi: 10.1016/j.neuroimage.2005.10.026

Apperly, I. A., & Butterfill, S. A. (2009). Do humans have two systems to track beliefs and belief-like states? *Psychological Review*, *116*, 953–970. doi: 10.1037/a0016923

Arbuthnott, K., & Frank, J. (2000). Trail Making Test, part B as a measure of executive control: Validation using a set-switching paradigm. *Journal of Clinical and Experimental Neuropsychology*, *22*, 518–528. doi: 10.1076/1380-3395(200008)22:4;1-0;FT518

Berg, E. A. (1948). A simple objective technique for measuring flexibility in thinking. *Journal of General Psychology*, *39*, 15–22.

Białek, A. (2011). Podejmowanie perspektywy epistemicznej przez osoby w różnym wieku [Epistemic perspective-taking in people of different age]. *Psychologia Społeczna*, *63*(18), 185–201.

Flavell, J. H., Everett, B. A., Croft, K., & Flavell, E. R. (1981). Young children's knowledge about visual perception – Further evidence for the Level 1–Level 2 distinction. *Developmental Psychology*, *17*, 99–103.

Forster, K. I., & Forster, K. C. (2003). DMDX: A Windows display program with millisecond accuracy. *Behavior Research Methods Instruments & Computers*, *35*(1), 116–124.

Happé, F. (1994). An advanced test of theory of mind: Understanding of story characters' thoughts and feelings by able autistic, mentally handicapped, and normal children and adults. *Journal of Autism and Developmental Disorders*, *24*, 129–154.

Hayes, A. F. (2018). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach* (2nd ed.). New York: Guilford Press.

[2] This relationship with the processing speed may be interpreted as reflecting the complex nature of the VPT-2 task. According to a simultaneity mechanism account (cf. Salthouse, 1996, p. 404) complex tasks require synchronization of the constituent operations, and synchronization is easier when the relevant operations can be executed rapidly.

Langdon, R., & Coltheart, M. (2001). Visual perspective-taking and schizotypy: Evidence for a stimulation-based account of mentalising in normal adults. *Cognition*, *82*, 1–26.

Michael, J., Wolf, T., Letesson, C., Butterfill, S., Skewes, J., & Hohwy, J. (2018). Seeing it both ways: Using a double-cuing task to investigate the role of spatial cuing in Level-1 visual perspective-taking. *Journal of Experimental Psychology: Human Perception and Performance*, *44*(5), 693–702. doi: 10.1037/xhp0000486

McDonald, L., & Stuart-Hamilton, I. (2003). Egocentrism in older adults: Piaget's three mountains task revisited. *Educational Gerontology*, *29*, 417–425. doi: 10.1080/713844355

Moll, G., & Tomasello, M. (2006). Level 1 perspective-taking at 24 months of age. *British Journal of Developmental Psychology*, *24*, 603–613. doi: 10.1348/026151005X55370

Pearson, A., Ropar, D., & Hamilton, A. F. (2013). A review of visual perspective taking in autism spectrum disorder. *Frontiers in Human Neuroscience*, *7*, 652. doi: 10.3389/fnhum.2013.00652

Perner, J., Brandl, J., & Garnham, A. (2003). What is a perspective problem? Developmental issues in understanding belief and dual identity. *Facta Philosophica*, *5*, 355–378.

Peterson, C. (2003). The social face of theory of mind: The development of concepts of emotion, desire, visual perspective, and false belief in deaf and hearing children. In B. Repacholi & V. Slaugther (Eds.), *Individual differences in theory of mind* (pp. 171–196). New York: Psychology Press.

Peterson, C., Peterson, J., & Webb, J. (2000). Factors influencing the development of theory of mind in blind children. *British Journal of Developmental Psychology*, *18*, 431–447. doi: 10.1348/026151000165788

Piaget, J., & Inhelder, B. (1948/1956). *The Child's Conception of Space.* New York: Norton.

Qureshi, A. W., Apperly, I. A., & Samson, D. (2010). Executive function is necessary for perspective selection, not Level-1 visual perspective calculation: Evidence from a dual-task study of adults. *Cognition*, *117*, 230–236. doi: 10.1016/j.cognition.2010.08.003

Reitan, R. M. (1971). Trail making test results for normal and brain-damaged children. *Perceptual and Motor Skills*, *33*(2), 575–581.

Salthouse, T. M. (1996). The processing-speed theory of adult age differences in cognition. *Psychological Review*, *103*(3), 403–428.

Samson, D., Apperly, I. A., Braithwaite, J. J., Andrews, B. J., & Scott, S. E. B. (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*, 1255–1266. doi: 10.1037/a0018729

Samson, D., Houthuys, S., & Humphreys, G. W. (2015). Self-perspective inhibition deficits cannot be explained by general executive control difficulties. *Cortex: A Journal Devoted to the Study of the Nervous System and Behavior*, *70*, 189–201. doi: 10.1016/j.cortex.2014.12.021

Santiesteban, I., Catmur, C., Hopkins, S. C., Bird, G., & Heyes, C. (2014). Avatars and arrows: Implicit mentalizing or domain-general processing? *Journal of Experimental Psychology: Human Perception and Performance*, *40*, 929–937. doi: 10.1037/a0035175

Sánchez-Cubillo, I., Periáñez, J. A., Adrover-Roig, D., Rodríguez-Sánchez, J. M., Ríos-Lago, M., Tirapu, J., & Barceló, F. (2009). Construct validity of the Trail Making Test: Role of task-switching, working memory, inhibition/interference control, and visuomotor abilities. *Journal of the International Neuropsychological Society*, *15*, 438–450. doi: 10.1017/S1355617709090626

Saxe, R., & Kanwisher, N. (2003). People thinking about thinking people: The role of the temporo-parietal junction in "theory of mind." *NeuroImage*, *19*, 1835–1842.

Surtees, A. D., Butterfill, S. A., & Apperly, I. A. (2012). Direct and indirect measures of Level-2 perspective-taking in children and adults. *British Journal of Developmental Psychology*, *30*, 75–86. doi: 10.1111/j.2044-835X.2011.02063.x

Wood, D. (2006). *How children think and learn: The social context of cognitive development*. Oxford, UK: Blackwell.

Zelazo, P. D., & Müller, U. (2002). Executive function in typical and atypical development. In U. Goswami (Ed.), *Handbook of childhood cognitive development* (pp. 445–469). Oxford, UK: Blackwell.