Other Papers

Polish Psychological Bulletin
2014, vol 45(4), 500-510
DOI - 10.2478/ppb-2014-0060

Adam Putko *
Agata Złotogórska **

Predictions of Actions and Their Justifications in False-Belief Tasks: The Role of Executive Function

Abstract: *The main objective of this study was to examine whether children's ability to justify their action predictions in terms of mental states is related, in a similar way as the ability to predict actions, to such aspects of executive function (EF) as executive control and working memory. An additional objective was to check whether the frequency of different types of justifications made by children in false-belief tasks is associated with aforementioned aspects of EF, as well as language. The study included 59 children aged 3–4 years. The ability to predict actions and to justify these predictions was measured with false-belief tasks. Luria's hand-game was used to assess executive control, and the Counting and Labelling dual-task was used to assess working memory capacity. Language development was controlled using an embedded syntax test. It was found that executive control was a significant predictor of the children's ability to justify their action predictions in terms of mental states, even when age and language were taken into account. Results also indicated a relationship between the type of justification in the false-belief task and language development. With the development of language children gradually cease to justify their action predictions in terms of current location, and they tend to construct irrelevant justifications before they begin to refer to beliefs. Data suggest that executive control, in contrast to language, is a factor which affects the development of the children's ability to justify their action predictions only in its later phase, during a shift from irrelevant to correct justifications.*
Key words: *preschoolers, false-belief task, theory-of-mind development*

Introduction

In studies on the development of theory of mind in children one of the most frequently used measures are false-belief tasks (e.g. Wimmer & Perner, 1983). The test question in these tasks usually requires a child to predict where the story protagonist will look for the desired object, which in his absence was moved to another location, or to determine the belief content of the protagonist about the location of the object. Numerous studies using false-belief tasks have shown that the majority of children begin to answer the prediction questions correctly around 4 years of age (see Wellman, Cross, & Watson, 2001). A less commonly used measure of the theory of mind is a question which requires a child to justify her previously made prediction regarding the protagonist's action, or to explain the presented action. Previous studies have shown that the accuracy of responses to prediction questions is related to several factors, including executive function[1] (EF) (e.g. Carlson & Moses, 2001; Frye, Zelazo, & Palfai, 1995) and language (cf. e.g. Milligan, Astington, & Dack, 2007). In contrast to the children's ability to predict behaviour in terms of mental states, little is known about factors associated with their ability to justify these predictions. Hence, the main objective of the present study was to compare the level of both kinds of abilities in preschool-aged children, and to check whether they are associated with the same aspects of EF. In the next section we will briefly review the research on the ability of children to justify their action predictions and to explain actions in terms of mental states, and then we will present the main objectives of the present study.

Justifications of Predictions and Explanation of Actions in False-Belief Tasks

Studies on how children justify their action prediction and how they explain other people's behaviour in

*** Institute of Psychology, A. Mickiewicz University; adam.putko@amu.edu.pl**
**** Institute of Psychology, A. Mickiewicz University**

[1] *Executive function* refers to a group of cognitive processes involved in flexible goal-directed behaviour, such as inhibition, set shifting, and working memory (e.g. Zelazo, Müller, Frye, & Marcovitch, 2003).

terms of mental states use a similar, although not identical, procedure. In the case of justifications children first answer the question which requires them to predict someone's action (or to determine someone's belief), and then are asked to justify their answer (e.g. "Why will he look for it in this place?"). Justification should be distinguished from explanation, which consists in the demonstration of someone's action (e.g. looking for the critical object in an empty place) and then asking to explain why someone will look in this place. Both procedures in certain conditions, namely, when children are asked to justify their correct predictions, allow for the assessment of the same ability – to explain behaviour. However, when action predictions are not correct, and children are asked to justify them, then we do not learn if they are able to explain actions in terms of mental states, but we gain an insight into how they understand the situation presented in the false-belief task. For this reason, it seems that it is the analysis of children's prediction justifications, not explanations of behaviour, that allows us to better understand the nature of theory-of-mind development. However, since both procedures usually produce similar results, they generally are discussed together in the literature.

The majority of studies comparing children's ability to predict actions and justify their predictions in terms of mental states showed an advantage of the former ability. For example, Clements and Perner (1994) found that 33% of children aged 2;5 to 4;6 years were able to make accurate predictions, whereas merely 16% could justify them correctly. Similar results were obtained by Moses and Flavell (1990) as well as Clements, Rustin, and McCallum (2000). In the latter study, correct justifications were made by 30% of those children (aged 2;10–5;0 years) who were able to make correct action predictions. However, in the study by Wimmer and Weichbold (1994), which used a mixed[2] procedure of justification/explanation, there was no significant difference between the frequency of correct predictions and justifications or explanations.

In the case of studies concerning children's ability to explain actions in terms of mental states the results are less consistent. Bartsch and Wellman (1989, exp. 2) found that correct answers to explanation questions were given by 66% of 3-year-olds, significantly more than in the case of questions which required making predictions of the story protagonist's action (31%) (p. 961). But as Perner and Lang (2000) have noted, in the group of children who gave a correct explanation the majority were those who responded only after a prompt question ("What does he think?"). Considering only those children who gave correct answers spontaneously, it turns out that the frequency of correct explanations did not differ significantly in Bartsch and Wellman's study from that of correct predictions.

Wimmer and Mayringer (1998), using the procedure of justifications/explanations, found no significant difference (when analysing data from full sample) between the frequency of correct predictions and explanations of actions. However, Wimmer and Gschaider (2000) reported that children had more difficulty with clearly explaining someone's wrong behaviour (searching in an empty place) than with predicting it. Robinson and Mitchell (1995), using the so-called twins task, which allows for the assessment of the ability to explain actions by forced-choice (rather than open-ended) questions, showed that explanation is easier than prediction. However, as Perner and Lang (2000) noted, when a correction for guessing was made for the observed frequency of correct answers in Robinson and Mitchell's study, the difference between the frequency of the proper predictions and explanations became very small. The issue of the difficulty of questions requiring the prediction and explanation of behaviour was also examined by Hughes, Dunn, and White (1998), who found a higher frequency of correct explanations as compared to predictions, and by Perner, Lang, and Kloo (2002), who found an advantage of prediction over explanation.

Summarizing the results of the few studies concerning children's ability to predict, justify or explain actions in terms of mental states, it should be noted that they varied, which probably could be attributed to differences in the types of questions and tasks used in particular studies ("deceptive box" or "unexpected transfer" false-belief tasks), which makes them not fully comparable. More conclusive results were obtained in studies comparing the frequency of correct predictions and their justifications, which showed that children have more difficulty making correct justifications than predictions. Therefore, it seems that the best way of assessing children's ability to predict actions and to justify these predictions is to measure both predictions and justifications within the same task. Such a solution was adopted in the present study, which aimed at comparing the level of both kinds of abilities in preschool-aged children.

Predictions and Explanations of Actions in Relation to Executive Function

So far there was little interest in factors related to the development of children's ability to justify or explain actions in terms of mental states. The only exception is a study by Perner et al. (2002, exp. 2), which was an attempt to examine whether children's ability to explain actions in false-belief tasks is associated with some aspects of EF. Two measures of EF were used in this study: the go-nogo task and the Dimensional Change Card-Sorting task (DCCS). It turned out that the DCCS task correlated more strongly with both measures of theory of mind than the go-nogo task. When age, verbal intelligence, and performance on the control questions in false-belief tasks were taken into account, only the relationship between the DCCS task and the explanation of behaviour remained significant. According to the authors of the study, a weaker relationship between the explanation of actions and performance on the go-nogo task, in comparison to the DCCS task, could be

[2] Both procedures were combined in such a way that if a child made a correct prediction, he was then asked to justify the answer. In the case of an incorrect prediction, the character's action for searching in an empty location was demonstrated, and the child had to explain it.

explained by the fact that the first task requires merely a simple ability to inhibit response, but not the type of executive control required by the DCCS task, i.e. the ability to refrain from prepotent, but incorrect response and to activate the opposite response instead.

Thus, the study of Perner and colleagues indicates that inhibitory demands are not responsible for the relationship between performance on the explanation task and the DCCS task, because the former is devoid of executive demands associated with inhibition of dominant responses, even though it correlates with executive control. However, Perner et. al.'s study does not allow for ruling out that the other component of executive control, i.e. working memory, is responsible for this correlation. According to Carlson, Moses, and Breton (2002), the relationship between performance on false beliefs tasks and "conflict" EF tasks – the measures of executive control – is best explained by a model assuming the combination of inhibition and working memory. They argue that both types of abilities are important for effective social cognition because it requires "children to simultaneously hold in mind two representations (working memory), and to select the representation and corresponding behavioural response that directly conflicts with children's own salient perspective, which must be actively suppressed (inhibition)" (p. 86). If, therefore, inhibition and working memory are essential components of executive control, and if the explanation of behaviour is devoid of inhibitory demands, then it cannot be ruled out that the factors responsible for the correlation between the explanation of behaviour and executive control are mainly the working memory demands or their interaction with inhibitory demands.

The role of working memory in the development and use of theory of mind was pointed out many times in the literature. Generally, there are two perspectives regarding the role of information processing capacity in the domain of theory of mind. According to the first one, an increase in cognitive resources allows a child to express or to apply theory-of-mind abilities (e.g. Leslie & Thaiss, 1992; cf. also Apperly, Samson, & Humpheys, 2009). However, according to the second view, an increase in cognitive resources is not only responsible for the expression, but also for the development (acquisition) of theory of mind (domain-general approach) (cf. e.g. Davis & Pratt, 1995; Gordon & Olson, 1998). Although subsequent longitudinal studies have shown that the role of working memory seems to be restricted to the factor which influences only the expression of theory of mind, the latter was assessed almost exclusively by tasks which did not require justification or explanation of actions. This raises the question about the role of working memory and inhibition skills in the development and use of the ability to justify or explain behaviour in terms of mental states.

The Aims of This Study

The present study had three objectives. The first was to examine whether there is a difference in children's ability to predict behaviour and to justify these predictions in terms of mental states. As noted above, current studies do not provide a definite answer on this issue. Since one of the possible factors responsible for the diversity of findings may be that both kinds of abilities were assessed by different tasks, in the present study they have been measured by questions asked within the same task, which allowed for better control of additional requirements posed by false-belief tasks.

The second objective of the present study was to investigate whether children's ability to justify their predictions in false-belief tasks is related to executive control and working memory, factors known to be associated with the level of theory of mind development in children when assessed with false-belief tasks (cf. e.g. Carlson et al., 2002; Gordon & Olson, 1988). Since an important factor associated with theory of mind is language (cf. Milligan et al., 2007), in the present study we decided to control its influence by measuring one of its aspects.

Third, and finally, the aim of the present study was to examine the relationship between two aspects of EF, executive control and working memory, as well as the frequency of different types of justifications made by children in false-belief tasks. Studies show that in the group of children providing incorrect justifications there is a large variation in responses. Their responses may refer to the current location of the searched object or to the protagonist's goals. They may also refer to irrelevant facts or give uninformative repetition of the answer (e.g. "right here"). Finally, the child may say that she does not know why someone will look for the critical object in this location, or may simply not give any answer.

As yet there was no interest in the question of whether the children's tendency to use certain types of justifications is associated with their level of cognitive development, including EF. Only Perner et al. (2002), as mentioned earlier, were interested in the relationship between children's ability to explain actions in false-belief tasks and EF, but the analysis of explanation answers was restricted to their correctness, without considering their types. Meanwhile, taking the types of children's justifications into account seems to be important, because changes in this respect may be an indicator of changes in children's understanding of mind, especially in those children who are able to make correct action predictions but still cannot justify them properly. The aim of the present study was, therefore, to take a more detailed look at this potential association by examining whether the child's tendency to give certain types of justifications is related to executive control and working memory.

METHOD

Participants

The final sample consisted of 59 children aged 3 to 4 years (34 girls and 25 boys) (*M* = 3;11 years, *SD* = 4.61 months, range 3;0–4;10), attending four nursery schools in one of the largest cities in Poland. In the group of 3-year-olds there were 31 children (*M* = 3;7 years, *SD* = 3.26 months, range 3;0–3;11), and in a group of 4-year-olds there were 28 children (M = 4;3 years, SD = 2.34 months, range 4;0–5;10). Additional three children participated in the study but due to them not completing some of the tasks (absence of a child during the second session or refusal to participate in the second session) their data were not included in the analysis. For all children participating in the study a written parental consent was obtained beforehand.

Materials and Procedure

Theory of Mind

Three false-belief tasks in an "unexpected transfer" version were used, and were presented in the form of stories enacted with miniature figures and props (Wimmer & Perner, 1983). In each task the story protagonist (a boy or a girl) put the critical object (an apple, a sandwich, or a sweet) in one of three locations, and then left the scene. At this point the story was stopped and the child was asked a control question concerning the location of the critical object. When the child answered, another character appeared on the scene (introduced as a brother or sister of the protagonist), discovered the critical object and ate it. After the protagonist returning, the child was asked several test and control questions. The first test question required the child to predict the protagonist's action: "Where will the boy first look for *p*?", and the second required him to justify this prediction: "Why will the boy look for *p* just there?" The control questions checked whether the child remembered all the relevant details of story: 1) "Where is p now?" 2) "Did the boy see that the girl ate p?" The child received one point for the correct response to the action prediction question if the correct response to this question was accompanied by correct responses to the control questions.

We used a version of false-belief tasks with reduced salience of reality, in which the critical object disappears (is eaten) instead of being simply transferred from one place to another. As a meta-analysis by Wellman et al. (2001) showed, the decrease of salience helps children – especially younger ones – to obtain better results in false-belief tasks, which was precisely the reason for using this version of task in the current study. Adding the word "first" to the test question was an additional factor aimed at preventing misinterpretation of this question by children.

Justification responses were coded by two raters on the basis of records in the protocols of the study. We used a coding scheme utilised by Clements and Perner (1994), which included the following categories of justifications: belief, past action, irrelevant justification, current location, no explanation, and goal (see Appendix). Cohen's kappa for the coding of justification answers was .89. Cases of disagreement were rectified by discussion.

Executive Control

In order to assess executive control Luria's hand-game was used (Luria, Pribram, & Homskaya, 1964; Hughes, 1998). Successful performance on this task requires a child to hold an arbitrary rule in working memory and inhibit a prepotent response in order to perform a rule-governed motor act. The rationale for the use of this task was that it appears to be a sensitive measure of executive control, especially among younger children (cf. Carlson, 2005; Flynn, O'Malley, & Wood, 2004). Following Hughes' (1998) procedure, two conditions were administered. In the imitation condition, children were asked to imitate the experimenter's hand movements (e.g., make a fist or open a hand). In the conflict condition, children were asked to execute the opposite movement to that of the experimenter; that is, when she made a fist, children had to open the hand, and when she opened her hand, children had to make a fist. Feedback was provided after each trial in the imitation condition and after two preliminary trials in the conflict condition. Fifteen test trials were presented in a pseudo-random order in the conflict condition with no more than two trials of the same type in a row. Children received 1 point for each correctly executed action in the conflict condition (out of 15), with high scores reflecting good executive control.

Working Memory

Counting and Labelling (Gordon & Olson, 1988). We also used a dual-task procedure, which required simultaneous execution of two actions – labelling pictures and enumerating. In the first part of the task the experimenter demonstrated required actions. First, she named and pointed to each picture in turn: "it's a monkey, it's a doll, it's a rabbit." Then, while pointing once again to pictures, she loudly counted: "one, two, three." Finally, she demonstrated the performance of both actions at the same time, pointing to the pictures and saying: "one is a monkey, two is a doll, three is a rabbit." In the second part of the task the child was instructed to repeat the steps the experimenter had performed. The child was corrected, as needed, after steps one and two, but not after step three. One point was given for each "name–number" pair specified correctly and in the proper order (to the maximum of 3). It is assumed that coordination of actions in the dual-task engages the central executive of working memory (Baddeley & Hitch, 1974; Baddeley, 2001), hence performance on this task by children can be regarded as a measure of working memory capacity. Using this task Gordon and Olson (1998) found that dual-tasks processing was significantly related to performance on false belief tasks (cf. also Carlson et al., 2002).

Language

To control for children's language skills we used an embedded syntax test, modelled on the study of Ruffman et al. (2003, Exp. 2). The test consisted of eight items in the form of sentences like: "The leaf that is under the ball is red." [Polish: "Liść, który jest pod piłką, jest czerwony"]. The items named two colours (red or green), two spatial relations (above or under), and 15 common objects (ball, leaf, crayon, bird, apple, cup, rabbit, spoon, car, cow, sun, dog, star, book and lamp). For each item the experimenter read a sentence aloud and asked the child to choose the correct picture from four possibilities. Eight test trials were preceded by checking whether the child correctly identifies the colours and distinguishes the relation "under–above". For every correct answer the child received 1 point (range

0–8). Reliability of the test was acceptable, Cronbach's alpha was .64.

General Procedure

Children were assessed individually by a female experimenter in a quiet room at kindergarten, in two sessions lasting approximately 15 minutes each. During the first session two false-belief tasks and the language test were administered, and during the second session the Counting and Labelling task, the last false-belief task and Luria's hand-game were administered. Children's responses were scored online during the assessment.

RESULTS

As shown in Table 1, most children responded to the action prediction questions in all three trials of false-belief tasks incorrectly or correctly. Assuming that to be categorized as having a good understanding of false beliefs a child had to give a correct answer at least in two trials out of three, this criterion for the action prediction questions was met by 55.9% of the children. With a more stringent criterion (all three trials correct) this percentage equalled 42.4%. The latter criterion, although reducing the probability of passing false-belief tasks by chance, seems to be too restrictive for the justification answers, because it seems unlikely that children would give three correct justifications by chance. For this reason in the analysis of performance on false-belief tasks we used a criterion of at least two correct (or identical, in the analysis of the frequency of types of justification) answers out of three. The analysis based on these data showed significant differences in the frequency of justification categories used by children (χ^2 (5, $N = 59$) = 38.40, $p < .001$). Children most frequently gave irrelevant justifications (39%), while they justified their answers in terms of belief or goal least frequently (1.7% in both cases). Considering the two categories of correct justifications – belief or past action – 28.8% of children gave at least two such justifications, which was a significantly lower percentage compared to those children, who gave also at least two correct predictions (55.9%, Wilcoxon Signed Ranks test $Z = 3.14$, $p = .001$).

Table 2 shows that 17 children gave at least two correct predictions in false-belief tasks, while at the same time not providing even one correct justification. The opposite pattern of results, correct (at least twice) justifications, accompanied at the same time by incorrect predictions, was shown by only two children. Therefore, comparison of the frequency of correct predictions and justifications (based on dichotomised data) indicates a significantly higher frequency of correct predictions (McNemar test, $p = .001$).

Distribution of scores from Luria's hand-game was strongly negatively skewed, with the median of 14 and dispersion ranging from 0 to 15 points. Because the attempts at data transformation did not improve their distribution, data from this task were dichotomised on the basis of the median. 67.8% of children obtained scores equal to or higher than the median. Distribution of scores on the Counting and Labelling task was close to bimodal: 19 children received 0 points, 3 children 1 point, 9 children 2 points, and 28 children 3 points. Here as well data were

Table 1. Scores for the Action Prediction and Justification Questions in False-Belief Tasks

Measure and category	Scores				% of Children Who Scored at Least 2 Points	% of Children Who Scored at Least 3 Points
	0	1	2	3		
Predictions						
	22	4	8	25	55.9	42.4
Justifications						
Belief	57	1	1	0	1.7	0
Past action	39	4	5	11	27.2	18.6
Irrelevant	27	9	6	17	39.0	28.8
No explanation	46	5	3	5	13.6	28.8
Current location	46	4	5	4	15.3	6.8
Goal	57	1	1	0	1.7	0

Table 2. Number of Children Who Made Correct Predictions in Relation to Correct Justifications in False-Belief Tasks

Correct Justifications	Correct Predictions				Total
	0	1	2	3	
0	19	2	7	10	38
1	1	1	0	2	4
2	1	0	1	3	5
3	1	1	0	10	12
Total	22	4	8	25	59

dichotomized on the basis of the median, which equalled 2 (62.7% of children received scores equal to or higher than this value). Only scores on language test did not differ from a normal distribution, $M = 4.31$, $SD = 2.00$, range 0–8. There was a weak relationship between performance on the Counting and Labelling task and Luria's hand-game ($\varphi = .294$, $p < .05$), reflecting the common requirements of both tasks, probably related to working memory.

Relating Predictions and Justifications to Executive Function and Language

To examine whether the children's ability to predict actions and justify these predictions was related to executive control, working memory, and language, an analysis of correlations was conducted. In all correlational analyses, in addition to the Pearson's r coefficient for the variables that met the assumptions of normal distribution, we used two other coefficients: Yule's phi coefficients of associations for dichotomised variables, and point-biserial coefficients for the relationships between the dichotomized and quantitative variables. As shown in Table 3, correct predictions in false-belief tasks were positively correlated with age and language, as well as with performance on the Counting and Labelling task. This last finding is consistent, among others, with the study by Gordon and Olson (1998). On the other hand, and contrary to expectations, the frequency of correct predictions did not significantly correlate with performance on Luria's hand-game. In the case of justification answers there was a positive relationship between them and age, language, and Luria's hand-game.

Results of the Logistic Regression Analysis

In further analysis we examined to what extent the relationship between correct predictions or justifications and the two aspects of EF is specific, i.e. independent of age and language. For this purpose, a series of hierarchical logistic regression analyses were conducted.

Prediction Answers

In the first analysis, the dependent variable was the frequency of correct action predictions in false-belief tasks, and the independent variable was performance on the Counting and Labelling task. Age as control variable was entered into the model in the first step, and the language ability in the second. The scores on the Counting and Labelling task were entered last. The analysis showed that a model containing three variables was well matched to the data (χ^2 (3, $N = 59$) = 9.43, $p = .024$, Nagelkerke's $R^2 = .198$). However, the Counting and Labelling task accounted for a non-significant (3.2%) increase in the variance of the dependent variable. Also, when the order of entered variables was changed to check whether the impact of language is specific, i.e. independent of age and performance on the Counting and Labelling task, the results were similar, with language responsible for a non-significant (3.0%) increase in the explained variance. As shown in Table 4, neither language nor performance on the Counting and Labelling task were significant predictors of action predictions when the effect of age was controlled[3].

Table 3. Relationships Between Correct Predictions and Justifications in False-Belief Tasks and Age, Executive Function, and Language ($N = 59$)

	Predictions	Justifications
Age	.310**	.399**
Luria's hand-game	.119	.358**
Count and Label	.233**	.104
Language	.257*	.449**

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

Table 4. Results of Hierarchical Logistics Regression Analysis for Action Predictions as a Dependent Variable

Step	*Nagelkerke's ΔR^2*	χ^2	*B*	*SE*	*p*	*Exp(B)*
Step 1	.128	5.92*				
Age			.149	.066	.023	1.161
Step 2	.038	1.88				
Language			.209	.155	.178	1.233
Step 3	.032	1.62				
Count and Label			.764	.602	.204	2.146

Note. *ΔR^2*- increment in variance accounted for; *B* -unstandardised regression coefficient.
*$p < .05$.

[3] The pattern of relationships changed, however, when the correctness of action predictions was determined on the basis of a more rigorous criterion of three correct answers, instead of just two. It turned out that correctness of action predictions does not correlate with performance on the Counting and Naming task, but does correlate with Luria's hand-game. The latter variable explained 12.1% of the variance in action predictions, independent of age and language (all variables explained 34.1% of the variance). With age and performance on Luria's hand-game controlled, language entered into the model in the second step accounted for 11.0% of the variance in action predictions.

Justifications Answers

In the next series of analyses the dependent variable was the frequency of correct justifications in false-belief tasks. As before, age and language were entered into the model as control variables to check whether the impact of executive control, measured with Luria's hand-game, is specific. The analysis showed that performance on Luria's task contributed a 17.4% increase in the explained variance when age and language development were taken into account (see Table 5). All three variables explained in total 54.2% of the variance in justification answers (χ^2 (3, N = 59) = 28.12, $p < .001$). Next, when language was entered into the model in the final step, with age and performance on Luria's hand-game controlled, it explained 7.1% ($p < .05$) of the variance of the dependent variable[4].

Summarising the results of the regression analysis, it should be noted that when age and language were statistically controlled, the only significant predictor of the children's justifications in false-belief tasks was executive control (taking a criterion of at least two correct responses out of three). Working memory, measured by the Counting and Labelling task, was not a significant predictor of either justifications or predictions, when age alone or both age and language were taken into account.

Types of Incorrect Justifications in Relation to Executive Function and Language

To examine whether the frequency with which the children referred to different types of erroneous justifications in false-belief tasks was related to EF and language, a correlation analysis was conducted in the group of 42 children who gave at least two wrong justification answers. Justification answers in terms of goal were excluded from the analysis due to the fact that only one child gave such answers at least twice. As shown in Table 6, none of the aspects of EF were significantly associated with a tendency to use a specific type of erroneous justification. Only language ability was significantly related to irrelevant justifications, and marginally significantly to other types of justifications.

Logistic regression analysis showed that language explained, independently of age, 22.5% of the variance in irrelevant justifications (B = .688, p = .012). To rule out that the lack of a significant correlation between a tendency to use various types of erroneous justifications and EF is the result of small variance of EF skills in this group of children, we checked what percentage of these children had high scores (equal to or higher than the median) on Luria's hand-game and on the Counting and Labelling task. It turned out that for this first task the percentage was 57.1% of children and for the latter it was 59.5%. As the proportion of these children was only slightly lower than in a full study sample, the lack of a significant relationship between the

Table 5. Results of Hierarchical Logistic Regression Analysis for Justifications as a Dependent Variable

Step	*Nagelkerke's* ΔR^2	χ^2	*B*	*SE*	*p*	*Exp(B)*
Step 1	.240	10.82***				
Age			.254	.091	.005	1.289
Step 2	.128	6.75**				
Language			.434	.185	.019	1.543
Step 3	.174	10.55**				
Count and Label			3.121	1.246	.012	22.674

Note. ΔR^2- increment in variance accounted for; *B* - unstandardised regression coefficient.
** $p < .01$; *** $p < .001$.

Table 6. Correlations Between Frequency of Different Types of Erroneous Justifications and Age, EF, and Language (n = 42)

	Justifications		
	Irrelevant	No explanation	Current location
Age	.257#	-.150	-.209
Luria's hand-game	-.014	-.193	.101
Count and Label	.128	-.029	-.160
Language	.422**	-.263#	-.279#

Note. ** $p < .01$; * $p < .05$; # $p < .1$.

[4] A similar analysis, carried out for the correctness of justifications in false-belief tasks determined on the basis of a more rigorous criterion of three correct answers, showed that executive control was responsible for 7.1% and language for 10.5% of the variance of dependent variable (all $p_s < .05$).

tendency to use various types of erroneous justifications and EF cannot be due to a small variance of EF skills in this group of children.

Differences Between Children Who Are Able to Justify Their Predictions and Those Who Are Able to Make Only Correct Predictions

In further analysis we examined whether children who were able to justify their predictions in false-belief tasks differed in terms of age, EF, and language from those children who were able to make only correct predictions. The first of these groups (n = 14) was slightly older (M = 49.5 months) than the other (n = 19, M = 47.6 months), but the difference was not statistically significant ($t_{(31)} = 1.40, p > .05$). The first group showed significantly higher scores on Luria's hand-game (Mann-Whitney's $U = 70, p = .021; r_g = .34$), and on the language test ($t_{(31)} = 3.11, p = .004$, Cohen's $d = 1.12$). There was no significant difference between these groups in performance on the Counting and Labelling task.

DISCUSSION

Predicting actions vs. justifying predictions in false-belief tasks

The study showed that significantly more children (about 29%) made correct predictions in false-belief tasks, while not being able to make correct justifications, as compared to children who showed the opposite pattern of results (3%). These proportions indicate the greater ease of predicting actions than justifying them in preschool-aged children. The results are consistent with other studies concerning children's ability to justify action predictions, such as Clements and Perner (1994) and Clements et al. (2000), as well as research which concentrated on the children's ability to explain actions in terms of mental states, rather than justify their predictions (e.g. Perner et al., 2002; Wimmer & Gscheider, 2000).

The question remains why children are able to make correct action predictions, taking into account someone's belief, earlier than they are able to justify them. In considering this issue, Clements and Perner (1994) drew attention to explicitation theory proposed by Karmiloff–Smith (1992), which assumes that many kinds of knowledge children acquire have initially an implicit form, to which children do not have conscious access. In the course of development this implicit knowledge is gradually transformed into an increasingly explicit form. At some level of development (E2) children gain conscious access to certain previously inaccessible content of acquired knowledge, and at the last (E3) level they become able to verbalise it. However, as noted by Clements and Perner (see also Clements et al., 2001), a problem for this theory in its present form is the developmental gap between the ability to predict and justify actions, as it seems that in 3- to 5-year-olds both these abilities are based on already verbalized knowledge, corresponding to the same – E3 – level, according to this theory. So without further assumptions or modifications Karmiloff–Smith's theory cannot explain the developmental gap between the children's ability to predict actions and to justify these predictions.

Another interpretation was put forward by Bartsch (1998). According to her, in order to explain someone's actions in terms of beliefs, children must first infer the state of someone's knowledge and then refer to that conclusion in their answer (stating, for example, that a story protagonist did not know where the object was placed). As Bartsch pointed out, it is easier for children to reiterate content already stated in and explicitly associated with the story than to verbally describe what was not explicitly connected, even if they understand what they wish to convey (p. 426). An additional problem is probably that children aged 3-4 years do not fully understand the concept of belief. Thus, according to Bartsch, an interaction of cognitive and language requirements may be responsible for children's difficulties in tasks involving an explanation of someone's behaviour.

The claim that not only language demands, but also conceptual ones, constitute the key factors in children's coping with justifying predictions or explaining actions in false-belief tasks, is not supported by other studies, for example, among others, Clements et al. (2000). In this study, children underwent training in solving false-beliefs tasks, consisting in explaining why a protagonist will look for the desired object in this but not in the other location. It turned out that the training had the greatest impact on the progress in justifying predictions among those children who could already make correct action predictions before training. As the authors noted, most of these children referred in their explanations to the concept of belief, although this term was not used during the training, when children were given justifications of correct predictions. Thus, the progress in the ability to justify predictions could not be the result of the repetition of justifications heard during the training, but probably was the result of conceptual change in understanding of false beliefs which has occurred in this group of children. This finding suggests that the ability to make correct predictions is a conceptual prerequisite for the development of a capacity to justify these predictions in terms of mental states. Some researchers (e.g. Bartsch, 1998) claim that the period in the development of theory of mind in which children make correct action predictions, but are not yet able to provide correct justifications (or explanations), is a transition period between an early mental theory of human actions and a more advanced one.

It seems that none of the explanations outlined above is comprehensive. In addition to language and new experiences, as well as confrontations of acquired knowledge with new critical events, resulting in the revision of existing knowledge, the data in the present study indicate an important role of yet another factor – executive control.

Executive Function and the Ability to Justify Action Predictions in False-Belief Tasks

The second objective of the present study was to examine whether children's ability to justify their action predictions is related to such aspects of EF as executive control and working memory. Relationships between these aspects of EF and the ability to predict actions in false-belief tasks were found in several previous studies (e.g. Carlson et al., 2002; Davis & Pratt, 1995; Gordon & Olson, 1998). The present study showed that executive control, as measured by Luria's hand-game, explained about 17% of the variance in the ability to justify predictions, even when influence of age and language were controlled for.

In contrast, working memory capacity, as measured by the Counting and Labelling task, was not associated with the ability to justify predictions in terms of mental states. The relationship between executive control and the children's ability to justify their action predictions is consistent with research by Perner et al. (2002), in which a similar relationship was found for the ability to explain actions. This indicates that both skills are similar to some degree. At the same time, lack of a significant relationship between working memory and justifications in false-belief tasks indicates that the factor responsible for the relationship between executive control and the children's ability to justify their predictions is not the working memory as such, but probably an interaction of working memory and inhibition, as suggested by Carlson et al. (2002), as well as the results of Perner et al.'s (2002) study, which ruled out the substantial role of inhibition as such.

It should be noted that, when compared with the results of analysis concerning the justifications of action predictions, rather surprising results were obtained for the other measure of theory of mind, predictions of actions. The frequency of correct predictions was not significantly associated with performance on Luria's hand-game, but with working memory capacity, as measured by the Counting and Labelling task. However, the relationship with the latter variable was no longer significant when age or language were taken into account. A potential factor responsible for the lack of a significant correlation between action predictions and executive control may be the type of false-belief task used in the present study. It was the task with reduced salience of reality, in which the critical object, instead of changing its location in the absence of the main story character, was portrayed as being eaten, so it "discontinued" to exist. This procedure usually results in an increase of correct answers (see meta-analysis by Wellman et al., 2001), which was the purpose of using this version of false-belief task in the present study. Reduced salience of reality involves, most likely, a decrease of executive requirements in false-belief task, consisting in overcoming the tendency to indicate the place of the actual location of the critical object. As far as we know, there was no research as yet on the link between EF and false-belief reasoning which used such kind of tasks, with reduced salience of reality. Most correlational studies used traditional false-belief tasks, not devoid of this factor. Thus, a weak relationship between performance on Luria's hand-game and the children's ability to make action predictions, found in the present study, could be the result of lower than usual executive requirements of false-belief tasks. This explanation is of course hypothetical and should be examined in an experimental study in which salience of reality would be a manipulated variable.

It should be noted, however, that a relationship between action predictions and executive control was revealed when the criterion of good performance on false-belief tasks was increased to three correct responses instead of only two. This finding may be interpreted in two ways. Firstly, if the tasks with reduced salience of reality pose lower executive requirements, then it may be difficult to capture the potential differences in false-belief reasoning, for which executive control would be responsible, in just two trials. However, these differences might become more noticeable if the number of trials was increased.

Alternatively, the higher criterion of passing false-belief tasks resulted in the removal from a group of children previously classified as understanding the false beliefs those who made correct predictions by chance. With two trials a probability to provide all correct responses by chance is not at all low, and is estimated at 25%, while with three trials it is estimated at 12.5%. The decrease of percentage of children who gave three correct action predictions to 42% (compared with 56% for two correct) may reflect the removal from this group of children of the so-called false positive cases. With an elevated criterion of passing false-belief tasks the assessment of false-belief understanding was probably more reliable, which revealed a significant relationship between correct action predictions and executive control, even when the influence of age and language was controlled for. It cannot be ruled out, however, that both aforementioned factors had contributed to the obtained pattern of results.

As regards the Counting and Labelling task, the present study showed that there is no significant relationship between correct justifications of action predictions in false-belief tasks and working memory capacity measured by this task. However, the latter variable was associated with the frequency of correct predictions. This last finding is consistent with Gordon and Olson (1998), who suggested that an increase in cognitive resources is not only responsible for the expression, but also for the development (acquisition) of theory of mind. It seems that the results of the present study do not support such a broad role of working memory. It is evidenced by the fact that there were no significant associations between working memory capacity and correctness of justifications or (as discussed below) the types of erroneous justifications (which would indicate, if such associations were found, the substantial role of working memory in the early phase of acquisition of theory of mind). Therefore, the role of working memory seems to be restricted to holding relevant information in mind while solving false-belief task, and as such it involves rather an expression of theory of mind than its development.

Relationships Between Language, EF, and Categories of Erroneous Justifications

The third objective of this study was to examine the potential associations between the categories of erroneous justifications in false-belief tasks, language, and EF. The study showed that, with regard to the development of language, children who are not able to make correct justifications tended to justify their action predictions less frequently in terms of "current location" (of searched object), and they also less frequently gave no explanation at all (marginally significant relationships). In contrast, the frequency of irrelevant justifications increased substantially with the development of language in this group of children and (marginally significantly) with age.

Interestingly, none of the aspects of EF were significantly associated with the frequency of various types of erroneous justifications. This pattern of correlations may suggest that language development fosters children's abandonment of justifications in terms of "current location" and leads them, before they become able to refer to beliefs, to increasingly use irrelevant justifications, consisting in uninformative repetition of responses (e.g., "right here") or referring to an irrelevant fact (e.g., "because this box is empty"). This could mean that with the development of language children gradually begin to discern the fallacy of their current thinking about causes of people's behaviour and begin to understand the need to justify their answers, although they cannot yet do that (a period of passing from a lack of justifications to irrelevant justifications). This pattern of results suggests that the category of irrelevant justification appears in theory-of-mind development later than other categories of erroneous justifications. Verification of this supposition would require longitudinal studies.

The fact that that there was no significant association between the frequency of different types of erroneous justifications and EF may indicate that executive control, in contrast to language, is not a factor which affects the development of the ability to justify action predictions from the very beginning, i.e. from the moment when children gradually cease to justify their action predictions in terms of the "current location" and shift to make irrelevant justifications. Its role begins to manifest only in the late period of development of justification competency, in a shift from irrelevant to correct justifications in terms of mental states. This is suggested by the latter type of conducted analysis, aimed at examining differences between children who were able not only to predict, but also to justify their predictions, and those who could only make correct predictions. It turned out that these two groups differ in terms of not only language development but also development of executive control.

CONCLUSIONS

The results of the present study lead to several important conclusions. Developing the ability to justify the action predictions in false-belief tasks is from its beginning associated with language development, which is generally consistent with the spirit of Karmiloff–Smith's (1992) theory of explicitation. The second important factor facilitating the development of justification competency appears to be executive control. As the justification of action prediction (like explaining someone's behaviour) is devoid of, as it is assumed, demands related to inhibition of prevalent but wrong responses, the link between the children's ability to justify predictions and executive control may indicate that this aspect of EF plays an important role not only in the expression of theory of mind but also in its development. The essence of this role is pointed out by executive emergence account (Moses, 2001; Moses, Carlson, & Sabbagh, 2005), according to which acquisition of belief concept "minimally requires some capacity to reflect on thought and action, some ability to distance oneself from the immediate situation, and some ability to inhibit salient but misleading knowledge" (Moses, 2001, p. 688).

Thus, the development of the ability to justify the action predictions seems to involve not as much the transformation of already existing knowledge about the mental causes of human behaviour into a more verbal form, as the construction of this knowledge by means of reflection on human behaviour. Such a role of executive control is also suggested by the results of Kloo and Perner's (2003) study, which found that training in solving the DCCS test, a measure of executive control, had a positive impact on theory-of-mind development. Unfortunately, the analysis presented by the authors of this study does not allow to determine whether training had a similar impact on the ability to predict actions and to justify this predictions, because both measures of theory of mind were combined in the analysis. Thus, in future studies it would be desirable to go beyond the paradigm of correlational studies and conduct more detailed analysis of the impact of training on different measures of false-belief understanding.

Another issue revealed by the present study is the influence of false-belief task in a version with reduced salience of reality on the strength of the relationship between the level of action predictions and executive control. Future work is needed to explore whether manipulation of this factor substantially influences the strength of the correlation between this measure of theory of mind and executive control.

Finally, some limitation of the present study is also a measure of language development, which focused on only one aspect of this development, that is, the syntax. Although research show that all aspects of language (semantics, grammar, and pragmatics) are related to the development of theory of mind (cf. e.g. Milligan et al., 2007), and also that the syntactic aspect is related to other aspects of language, the results obtained in the current study should be verified using other measures of language development.

References

Apperly, I.A., Samson, D., & Humpheys, G.W. (2009). Studies of adults can inform accounts of theory of mind development. *Developmental Psychology, 45*(1), 190-201.

Baddeley, A. (2001). Is working memory still working? *American Psychologist, 56*, 851-64.

Baddeley, A., Hitch, G.J. (1974). Working memory. In G.A. Bower (ed.), *The psychology of learning and motivation* (vol. 8, pp. 47-89). New York: Academic Press.

Bartsch, K. (1998). False belief prediction and explanation: Which develops first and why it matters. *International Journal of Behavioral Development, 22*, 423–428.

Bartsch, K., & Wellman, H.M. (1989). Young children's attribution of action to beliefs and desires. *Child Development, 60*, 946-964.

Carlson, S.M. (2005). Developmentally sensitive measures of executive function in preschool children. *Developmental Neuropsychology, 28*, 595-616.

Carlson, S.M., & Moses L.J. (2001). Individual differences in inhibitory control and children's theory of mind. *Child Development, 72*, 1032-1053.

Carlson, S.M., Moses, L.J., & Breton, C. (2002). How specific is the relation between executive function and theory of mind? Contributions of inhibitory control and working memory. *Infant and Child Development, 11*, 73-92.
Clements, W.A., & Perner, J. (1994). Implicit understanding of false belief. *Cognitive Development, 9*, 377-395.
Clements, W.A., Rustin, L.C., & McCallum, S. (2000). Promoting the transition from implicit to explicit understanding: A training study of false belief. *Developmental Science, 3*, 81-92.
Davis, H.L., & Pratt, C. (1995). The development of children's theory of mind: The working memory explanation. *Australian Journal of Psychology, 47*, 25-31.
Flynn, E., O'Malley, C., & Wood, D. (2004). A longitudinal, microgenetic study of the emergence of false belief understanding and inhibition skills. *Developmental Science, 7,* 103-115.
Frye, D., Zelazo, P.D., & Palfai, T. (1995). Theory of mind and rule-based reasoning. *Cognitive Development, 10*, 483-527.
Gordon, A.C.L., & Olson, D.R. (1998). The relation between acquisition of a theory of mind and the capacity to hold in mind. *Journal of Experimental Child Psychology, 68*, 70-83.
Hughes, C. (1998). Finding your marbles: Does preschoolers' strategic behaviour predict later understanding of mind? *Developmental Psychology, 34*, 1326-1339.
Hughes, C., Dunn, J., White, A. (1998). Trick or treat?: Uneven understanding of mind and emotion and executive dysfunction in 'hard to manage' preschoolers. *Journal of Child Psychology and Psychiatry, 39*, 981-994.
Karmiloff-Smith, A. (1992). *Beyond modularity. A developmental perspective on cognitive science*. Cambridge: Cambridge Univ. Press.
Kloo, D., & Perner, J. (2003). Training transfer between card sorting and false belief understanding: Helping children apply conflicting descriptions. *Child Development, 74*, 1823-1839.
Leslie, A.M., & Thaiss, L. (1992). Domain specificity in conceptual development: Neuropsychological evidence from autism. *Cognition, 43*, 225-251.
Luria, A.R., Pribram, K., & Homskaya, E. (1964). An experimental analysis of the behavioural disturbance produced by a left frontal arachnoidal endotheliomn (meningioma). *Neuropsychologia, 2*, 257-280.
Milligan, K., Astington, J.W., & Dack, L.A. (2007). Language and theory of mind: Meta-analysis of the relation between language ability and false-belief understanding. *Child Development, 78*(2), 622 – 646.
Moses, L.J. (2001). Executive accounts of theory-of-mind development. *Child Development, 72*, 688-690.
Moses, L.J., Carlson, S.M., & Sabbagh, M.A. (2005). On the specificity of the relation between executive function and children's theories of mind. In W. Schneider, R. Schumann-Hengsteler, & B. Sodian (eds.), *Young children's cognitive development. Interrelationships among executive functioning, working memory, verbal ability, and theory of mind* (pp. 131-146). Mahwah, NJ: Erlbaum.
Moses, L.J., Flavell, J.H. (1990). Inferring false beliefs from actions and reactions. *Child Development, 61*, 929–945.
Perner, J., & Lang, B. (2000). Theory of mind and executive function: Is there a developmental relationship? In S. Baron-Cohen, H. Tager-Flusberg, D.J. Cohen (eds.), *Understanding other minds: Perspectives from developmental cognitive neuroscience* (2nd ed., pp. 150-181). Oxford: Oxford Univ. Press.
Perner, J., Lang, B., & Kloo, D. (2002). Theory of mind and self-control: More than a common problem of inhibition. *Child Development, 73*, 752–767.
Robinson, E.J., & Mitchell, P. (1995). Masking of children's early understanding of the representational mind: Backwards explanation versus prediction. *Child Development, 66*, 1022-1039.
Ruffman, T., Slade, L., Rowlandson, K., Rumsey, C., & Garnham, W.A. (2003). How language relates to belief, desire, and emotion understanding. *Cognitive Development, 18*, 139-158.
Wellman, H.M., Cross, D., & Watson, J. (2001). Meta-analysis of theory-of-mind development: The truth about false belief. *Child Development, 72*, 655-684.
Wimmer, H., & Gschaider, A. (2000). Children's understanding of belief: Why it is important to understand what happened. In P. Mitchell & K. Riggs (red.), *Children's reasoning and the mind* (pp. 253-266). Hove, UK: Psychology Press.
Wimmer, H., & Mayringer, H. (1998). False belief understanding in young children: Explanations do not develop before predictions. *International Journal of Behavioral Development, 22*, 403-422.
Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition, 13*, 103-128.
Wimmer, H., & Weichbold, V. (1994). Children's theory of mind: Fodor's heuristics examined. *Cognition, 53*, 45-57.
Zelazo, P.D., Müller, U., Frye, D., & Marcovitch, S. (2003). The development of executive function in early childhood. *Monographs of the Society for Research in Child Development, 68*(3), Serial No. 274.

Appendix

Types of justifications and examples of answers in the false-belief task (coding scheme modelled on Clements and Perner, 1994)

Type of justification	Example
Belief	"Because he thought it was there."
Past action	"Because he put it there."
Irrelevant justification	"Because so"; "Because he lost it."
No explanation	"I don't know."
Current location	"Because it is there."
Goal	"Because he wants to eat it."