Journal of Experimental Child Psychology xxx (2008) xxx–xxx

Contents lists available at ScienceDirect

Journal of Experimental Child Psychology

journal homepage: www.elsevier.com/locate/jecp

Persistent overconfidence despite practice: The role of task experience in preschoolers' recall predictions

Amanda R. Lipko [*,1], John Dunlosky, William E. Merriman

Department of Psychology, Kent State University, Kent, OH 44242, USA

ARTICLE INFO

Article history:
Received 18 April 2008
Revised 10 October 2008
Available online xxxx

Keywords:
Metamemory
Memory
Metacognition

ABSTRACT

In three experiments, preschoolers' ability to predict their picture recall was examined. Children studied 10 pictures, predicted how many they would recall, and then attempted to recall them. This study–prediction–recall trial was repeated multiple times with new pictures on each trial. In Experiment 1, children were overconfident on the initial trial, and this overconfidence persisted across three trials. In Experiment 2, children predicted either their own performance or another child's performance. Their predictions were overconfident across all trials regardless of whether they made predictions for themselves or for another child, suggesting that wishful thinking cannot fully account for their overconfidence. In Experiment 3, some children postdicted their previous recall performance prior to making each prediction. Although their postdictions were quite accurate, their predictions were still overconfident across five trials. Preschoolers' overconfidence was remarkably resistant to the repeated experience of recalling fewer pictures than the children had predicted. Even asking them to report the number that they recalled on a previous trial, which they could do accurately, did not cause them to lower their predictions across trials.

Introduction

Young children are overconfident about their abilities. If you spend a day around preschoolers, you are likely to hear them claim that they can accomplish feats that are well above their actual capabilities. Such overconfidence is evident across a variety of physical tasks (e.g., Plumert, 1995;

* Corresponding author.
E-mail addresses: alipko@brockport.edu, alipko@kent.edu (A.R. Lipko).
[1] Present address: College at Brockport, State University of New York, Brockport, NY 14420, USA.

0022-0965/$ - see front matter © 2008 Elsevier Inc. All rights reserved.
doi:10.1016/j.jecp.2008.10.002

Plumert & Schwebel, 1997; Schneider, 1998; Stipek, Roberts, & Sanborn, 1984) and cognitive tasks (e.g., Cunningham & Weaver, 1989; Flavell, Friedrichs, & Hoyt, 1970; Schneider, 1998; Shin, Bjorklund, & Beck, 2007; Yussen & Levy, 1975). Although overconfidence declines with age (Schneider & Pressley, 1997), grade school students and even adults still show overconfidence in their memory performance (e.g., Lipko, Dunlosky, Rawson, Swan, & Cook, 2007; Schneider, Visé, Lockl, & Nelson, 2000), which in turn can lead to poor self-regulation and performance (Thiede, 1999). Unfortunately, the reasons why children are so overconfident during the preschool years are still not fully understood.

A major goal of the current investigation was to investigate preschoolers' overconfidence. More specifically, we assessed whether preschoolers' overconfidence would persist across multiple trials of a memory task or whether the children would incorporate task experience into their predictions and, thus, show less overconfidence across trials. In the remainder of the introduction, we review some of the previous research on children's overconfidence and then present the rationale for the current experiments.

Early research investigating children's overconfidence in their learning

In a classic study, Flavell and colleagues (1970) used a performance prediction paradigm in which nursery schoolers, kindergarteners, second-graders, and fourth-graders were asked to predict their memory span. On each trial, a single picture was added to the memory set (up to 10 pictures total) and children were asked whether they would be able to remember the pictures once they were removed. Children's predicted memory span was defined as the longest series of pictures they said they would be able to remember. After completing the prediction task, children's actual memory span was assessed with an aural recall task. An experimenter read a series of familiar object names, and then the children were asked to repeat the names in the correct order. Children's actual memory span was defined as the longest series of words they were able to recall correctly. In all age groups, children were overconfident in that their predicted span was higher than their actual span. Using a similar procedure, Yussen and Levy (1975) also found that preschoolers and third-graders overpredicted their performance.

More recently, Shin and colleagues (2007) asked kindergarteners to make predictions of their memory using a different methodology. Children were asked to predict how many of 15 pictures they would be able to remember. After making their predictions, they were shown 15 pictures and were asked to name each one as it was presented. The children were then given 2 min to study the pictures. At the end of the 2-min period, they completed a distractor task and attempted to recall the pictures they had just studied. Children's prediction values were significantly higher than their recall performance.

Although studies like the ones described above have demonstrated repeatedly that young children are overconfident, the degree of children's overconfidence across trials of a task has not been explored systematically. Investigating change in children's predictions across trials allows better evaluation of potential explanations of children's overconfidence. For example, the wishful thinking hypothesis posits that children's predictions are based on how they want to perform rather than on how they expect to perform (Stipek et al., 1984). If children base their predictions on wishful thinking, their predictions may remain equally overconfident across all trials. Another possibility is that children are merely unfamiliar with the task and begin by setting an optimistic anchor for their performance. If so, then with further task experience that involves recall attempts, their overconfidence may be reduced dramatically.

Rationale for the current experiments

In Experiment 1, children were asked to predict their picture recall. They named 10 pictures and, following a brief study period, predicted how many they would be able to recall. The pictures were covered, and the children attempted recall. They were then told how many pictures they had recalled, and a new trial began using 10 new pictures. They completed three of these study–prediction–recall trials, with new pictures being presented on each trial.

Previous studies that have reported trial-by-trial performance provide a few possibilities for the outcomes of this experiment. As per the wishful thinking hypothesis, one possibility is that children's predictions will remain equally overconfident on all three trials of the task. Consistent with this prediction, Shin and colleagues (2007) reported that kindergarteners remained equally overconfident across all five trials of a memory task. However, important methodological differences exist between the current experiments and this previous study. In the current experiments, children were asked to predict their performance *after* studying all of the pictures and while the pictures were still in view, whereas Shin and colleagues asked children to make their predictions *prior to* seeing the pictures or naming them. In the latter situation, children's overconfidence may have persisted across trials because they were unaware of the pictures they would be studying and trying to recall on each trial. Their predictions could have been based on their beliefs about their memory in general rather than about their ability to remember the particular items.

Another possible outcome of the current experiments is that children's predictions will be overconfident on all three trials, but the prediction values will decline across trials. Cunningham and Weaver (1989) found this pattern of results when they asked kindergarteners to predict their recall of aurally presented words. As in Shin and colleagues' (2007) study, their task also required children to make predictions about stimuli that had not yet been presented. Without experience with the particular to-be-learned items, people's judgments are based largely on their general beliefs about their memory and cannot be influenced by their monitoring of learning per se (Dunlosky & Hertzog, 2000; Hertzog, Dixon, & Hultsch, 1990). In general, research with college students and older adults indicates that predictions made prior to study are based on different factors from those made immediately after study, with the latter typically being more accurate (Hertzog, Saylor, Fleece, & Dixon, 1994).

Thus, in the current research, we evaluated whether persistent overconfidence occurs when preschoolers have an opportunity to study the list of to-be-recalled items prior to predicting their performance. In particular, the first goal of Experiment 1 was to estimate the accuracy of children's poststudy predictions of picture recall. The second goal was to evaluate whether their overconfidence could be eliminated with repeated recall experience. To foreshadow, children were overconfident across all of the trials. As explained in detail later, Experiments 2 and 3 were designed to evaluate potential explanations for why children's predictions remain overconfident despite practice. More specifically, Experiment 2 evaluated a wishful thinking explanation and Experiment 3 evaluated the degree to which inaccurate postdictions of previous performance limit the accuracy of children's subsequent predictions.

Experiment 1

Method

Participants

A total of 21 children (mean age = 5 years 0 months, range = 4 years 0 months to 5 years 11 months, 12 girls and 9 boys) were recruited to participate. Children were recruited through their preschools, which were located in a primarily suburban, middle-class section of northeastern Ohio.

Materials and procedure

Children were tested individually in a quiet room in their preschool. Children sat at a table next to the experimenter in front of a large white foam board that contained 10 magnets. To ensure that the children understood how many pictures there were, before beginning the task they were asked to count to 10 and count the number of magnets on the board in front of them. They were then asked a few practice questions to ensure that they could compare quantities of small groups of objects. After the practice questions, the experimenter explained the task to the children.

The task began with the experimenter placing 10 4 $\times$ 6-inch colored pictures of familiar objects on the magnetic board (the lists of pictures were always presented in the order listed in Appendix A). The pictures were taken from *My First Word Book* (Wilkes, 1999). Pilot testing determined that the pictures were highly familiar to preschool-age children. As each picture was placed on the board, the children were asked to name it. After all 10 pictures had been presented, the children were given 10 s to study the pictures (this presentation rate was chosen based on initial pilot testing so as to ensure that levels of recall were off the performance floor and ceiling). They were then asked, "How many of these pictures do you think you are going to be able to remember once I cover them up?" The pictures were covered with a large piece of red paper, and the children were asked to recall as many of the pictures as possible. After 60 s, the children were asked to stop trying to recall pictures and the pictures were uncovered. Most children stopped recalling pictures well before this time limit.

After uncovering the pictures, the experimenter told the children the number of pictures they had recalled. The experimenter also removed the same number of pictures from the board as a visual demonstration of this number. The entire procedure was repeated twice more, each time with 10 different pictures. At the end of the third trial, the children were given a small gift of stickers to thank them for their participation.

Results and discussion

Although gender is not a focus of the current research (nor has it been for previous research), we first conducted all analyses with gender as a factor. The main effect of gender and all interactions including it were not significant, so we do not discuss gender further.

Prediction accuracy

Mean prediction accuracy (predicted recall–actual recall) was first computed for each participant, and the means across participants' values are presented in Table 1. Higher values indicate greater overconfidence. A trend occurred toward greater overconfidence with practice, although the analysis of variance (ANOVA) indicated that this trend only approached statistical significance, $F(2, 40) = 3.20$, $p = .05$. Thus, even though the children had the opportunity to study the pictures before making their predictions, they behaved similarly to preschoolers in previous studies who made predictions before viewing the study items (Cunningham & Weaver, 1989; Shin et al., 2007).

Mean prediction and recall performance

Because prediction accuracy is derived from prediction and recall magnitudes, we also present these values for completeness (Table 2). Across all trials and participants, intrusion errors were minimal (only 3) and so were not included in these analyses. A 2 (Task: prediction or recall) × 3 (Trial: 1, 2, or 3) ANOVA yielded significant main effects of task, $F(1, 20) = 54.85$, $p < .01$, $\eta^2 = .73$, and trial, $F(2, 40) = 18.11$, $p < .01$, $\eta^2 = .47$, but no significant interaction. On all three trials, children's predicted recall ($M = 7.89$) was substantially higher than their actual recall ($M = 4.27$).

Table 1
Predictive accuracy (predicted recall – actual recall) across trials in Experiments 1, 2, and 3

	Trial 1	Trial 2	Trial 3	Trial 4	Trial 5
Experiment 1	2.81 (0.59)	4.52 (0.53)	3.52 (0.74)	–[a]	–[a]
Experiment 2					
Self condition	2.68 (0.43)	3.66 (0.52)	3.72 (0.53)	3.62 (0.52)	3.06 (0.55)
Other condition	2.62 (0.58)	1.69 (0.47)	1.97 (0.44)	2.66 (0.52)	2.56 (0.50)
Experiment 3					
Control group	3.93 (0.49)	4.39 (0.62)	3.81 (0.49)	4.19 (0.61)	4.03 (0.56)
Postdiction group	3.82 (0.53)	4.16 (0.52)	3.88 (0.58)	4.48 (0.57)	4.12 (0.70)

Note. Standard errors of the means are in parentheses. All values were reliably greater than zero, $p < .05$.
[a] Only three trials occurred in Experiment 1.

Table 2
Mean prediction, recall, and postdiction values in Experiments 1, 2, and 3

	Trial 1	Trial 2	Trial 3	Trial 4	Trial 5
Experiment 1					
Prediction	8.33 (2.37)	8.29 (2.22)	7.05 (2.73)	–[a]	–[a]
Recall	5.52 (1.60)	3.76 (1.34)	3.52 (1.50)	–	–
Experiment 2					
Self condition					
Prediction	6.75 (2.68)	6.25 (2.54)	6.34 (2.91)	6.16 (3.06)	5.44 (2.72)
Recall	4.06 (1.64)	2.59 (1.32)	2.59 (1.36)	2.44 (1.48)	2.37 (1.38)
Other condition					
Prediction	6.62 (3.27)	5.69 (2.66)	5.97 (2.47)	5.66 (2.94)	5.56 (2.83)
Recall[b]	4.00	4.00	4.00	3.00	3.00
Experiment 3					
Control group					
Prediction	8.90 (2.26)	7.71 (3.05)	7.03 (2.83)	7.19 (3.05)	7.22 (2.81)
Recall	4.97 (1.45)	3.32 (1.60)	3.22 (1.36)	3.00 (1.09)	3.19 (1.78)
Postdiction group					
Prediction	8.73 (2.45)	7.87 (2.70)	7.12 (3.21)	7.45 (3.09)	7.18 (3.44)
Recall	4.90 (1.66)	3.67 (1.38)	3.24 (1.41)	2.97 (1.16)	3.06 (2.56)
Postdiction	–[c]	5.90 (2.26)	4.61 (2.29)	3.87 (1.90)	3.76 (2.15)

Note. Standard deviations are in parentheses.
[a] Only three trials occurred in Experiment 1.
[b] Values represent the predetermined recall of the child model in the video.
[c] Postdictions were not collected on the first trial.

Concerning the main effect of trial, both prediction and recall values declined across the three trials. Predictions on the first trial were significantly greater than those on the third trial, $t(20) = 2.26$, $p < .05$, $d = .50$. The mean prediction on the second trial was also significantly different from that on the third trial, $t(20) = 2.62$, $p < .05$, $d = .50$. The mean recall on the first trial was significantly higher than that on the third trial, $t(20) = 5.80$, $p < .01$, $d = .93$, and the mean recall on the first trial was also significantly higher than that on the second trial, $t(20) = 4.75$, $p < .01$, $d = .79$. The recall difference between the second and third trials was not statistically significant, $t < 1$. This slight decline in recall performance could have any number of causes (e.g., interference, fatigue).

Correlations between predictions and recall

Correlations between predictions and recall (with age partialled out) are presented in Table 3. These correlations provide further insight into why children demonstrated persistent overconfidence. Most relevant are the correlations between recall on one trial and predictions on the *next* trial, which may indicate the degree to which the children used recall outcomes to adjust predictions. As shown in Table 3, children's predicted recall for a trial did not correlate significantly with the number they had

Table 3
Correlations between predictions and recall with age partialled out in Experiment 1

	P2	P3	R1	R2	R3
Prediction 1	.61**	.52	.08	.15	−.13
Prediction 2		.63**	**.09**	.16	.12
Prediction 3			.28	**.06**	−.24
Recall 1				.32	.48*
Recall 2					.35

Note. P2, prediction on Trial 2; P3, prediction on Trial 3; R1, recall on Trial 1; R2, recall on Trial 2; R3, recall on Trial 3. Bold values are the correlations between recall on Trial *N* and predictions on Trial *N* + 1.
* $p < .05$.
** $p < .01$.

Please cite this article in press as: Lipko, A. R., et al. Persistent overconfidence despite practice: The role ... *Journal of Experimental Child Psychology* (2008), doi:10.1016/j.jecp.2008.10.002

just recalled on the previous trial. This lack of relationship cannot be due to low reliability in either measure because stable individual differences were evident in predicted recall (median cross-trial $r = .62$) and in actual recall (median cross-trial $r = .35$).

Although less relevant to our current concerns, other correlations in Table 3 are informative to the larger literature on children's metamemory performance (Schneider & Lockl, 2008). In particular, correlations between predictions and recall on a given trial provide a between-participants measure of judgment accuracy. For the current preschoolers, these correlations were consistently low and not statistically different from zero.

Experiment 2

In Experiment 1, preschoolers' recall predictions were overconfident across all three recall trials. This pattern of persistent overconfidence may be explained by the wishful thinking hypothesis (Stipek et al., 1984). According to this hypothesis, children base predictions on how they desire to perform rather than on how they think they will actually perform. The goal of the second experiment was to evaluate this hypothesis.

Previous research has evaluated the wishful thinking hypothesis for both psychomotor tasks and cognitive tasks. Stipek and colleagues (1984) asked 4-year-olds to predict either their own performance (self-predictions) or another child's performance (other-predictions) on a task in which they used a pulley to move a cart to the top of a tower. Under one condition, children predicted how far up the tower they could move the cart without losing a marble that was balanced on the cart. In another condition, they predicted how far someone else could move the cart up the tower. If the wishful thinking hypothesis explains children's overconfidence, self-predictions should be significantly greater than other-predictions because preschoolers tend to desire greater levels of performance for themselves than for others (Stipek & Hoffman, 1980). Self-predictions did exceed other-predictions, supporting the wishful thinking hypothesis.

Schneider (1998, Experiment 1) replicated Stipek and colleagues' (1984) findings with a different psychomotor task in which 4- and 6-year-olds predicted how many balls they would shoot into a basket. Children's predictions for their performance exceeded their predictions for another child's performance. In further support of the wishful thinking hypothesis, the number of shots that one group of children predicted they would make on each trial was similar in magnitude to the number that another group of children indicated they *wished to make* on each trial.

Schneider (1998, Experiment 2) also evaluated the wishful thinking hypothesis with two memory tasks. In one task, children were shown 10 pairs of pictures, each consisting of two identical individuals. Children watched as the experimenter hid one picture from each pair in a different house. They were then asked to place each remaining picture above the location where its matching picture had been hidden. Prior to engaging in the task, children were asked to predict how many of the pictures they would be able to match correctly. In support of the wishful thinking hypothesis, the mean number of pictures that children indicated they expected themselves to match correctly on each trial was no different from the mean number that other children indicated they wished to match correctly on each trial. However, contrary to the wishful thinking hypothesis, self- and other-predictions for this task did not differ. It is not surprising that performance on the hide-and-seek task did not support the wishful thinking hypothesis. As Schneider (1998) discussed, the outcome of this hide-and-seek task was not perceptually salient for observers. Thus, children may have had more difficulty in making predictions for others.

In a second memory task, children stopped a tape recorder after they had heard as many words as they believed they could recall. In support of the wishful thinking hypothesis, the mean number of words that children indicated they *expected to* remember on each trial was no different from the mean number that other children indicated they *wished to* remember on each trial. Although the wishful thinking hypothesis may explain children's overconfidence on concurrent memory tasks such as this one, it might not fully account for the greater level of overconfidence that preschoolers are likely to show in a prospective memory task. Cunningham and Weaver (1989) found that preschoolers' predictions were more overconfident on a prospective task than on a concurrent memory span task.

Because the memory tasks used by Schneider (1998) may have minimized the contribution of wishful thinking to overconfidence, in Experiment 2 we evaluated this hypothesis with the memory task used in Experiment 1. Using the method employed by both Stipek and colleagues (1984) and Schneider (1998), we compared children's predictions for their own performance with their predictions for another child's performance. If wishful thinking is an explanation for children's substantial overconfidence on the current multitrial task, (a) their self-predictions will be significantly greater than their other-predictions (Schneider, 1998; Stipek et al., 1984) and (b) overconfidence will be consistently higher for self-predictions than for other-predictions.

Finally, although children's predictions declined across trials in Experiment 1, this decline was relatively minimal. Perhaps with further task experience, the rate of decline will increase and ultimately eliminate overconfidence. Accordingly, in Experiment 2, children participated in five study–prediction–test trials.

Method

Participants

A total of 32 children (mean age = 4 years 7 months, range = 4 year 0 months to 5 years 4 months, 20 girls and 12 boys) were recruited to participate. None of the children had participated in the first experiment. They were recruited through their preschools, which were located in a primarily suburban middle-class section of northeastern Ohio.

Design, materials, and procedure

A 2 (Prediction: self or other) × 2 (Order: self-prediction first or self-prediction second) × 2 (Video Model: female or male) mixed-design ANOVA was used. The first factor was manipulated within each participant. All other factors were counterbalanced and manipulated between participants. Thus, all children completed both prediction conditions (self and other), but the presentation order of the conditions was counterbalanced. Half of the children completed the self-prediction condition first, and the other half completed the other-prediction condition first. The sex of the model child in the other-prediction video was also controlled. Half of the male participants watched a female model, and half of the female participants watched a male model. Within this counterbalancing, presentation order of the conditions was controlled as well.

Children were tested individually in a quiet room in their preschool. Before beginning the task, they were asked to count to 10 and count the number of magnets on the board in front of them. They were then asked a few practice questions to ensure that they could compare quantities of small groups of objects. After the practice questions, the experimenter explained the task to the children. Children participated in two conditions of the task: a self-prediction condition and an other-prediction condition. The same pictures were used in both conditions, and these pictures were presented in the same order. After completing both conditions, the children were given a small reward of stickers.

Self-prediction condition. The materials and procedure were identical to those used in the first experiment with one exception: Two additional trials, each with 10 new pictures, were added to the task for a total of five trials (see Appendix A for lists of pictures).

Other-prediction condition. Children were told that they were going to watch a DVD of a boy or girl of their age playing a memory game. On the DVD, a male or female preschooler (the model) performed five trials of the self-prediction task with the experimenter (the first author). Each trial began by asking the participant to watch the opening scene of the DVD in which the model looked at 10 pictures (same as those used in the self-prediction condition) on the magnetic board. The experimenter paused the video, pointed to and named each of the pictures on the screen for the participant, and asked, "How many pictures do you think Adam [Morgan] is going to be able to remember once the pictures are covered up?" After the prediction was given, the video resumed and the participant watched the model attempt to recall the pictures. The model was instructed to recall predetermined pictures. Over the five trials, the model recalled 4, 4, 4, 3, and 3 pictures, respectively. This pattern of recall was chosen to reflect the average level of recall found across trials in Experiment 1. After observing the model's recall, the experimenter told the participant how many pictures the model (Adam or Morgan) had been able to recall and then the next trial began.

Results and discussion

All analyses were first conducted with gender as a factor, but the main effect of children's gender and all interactions including it were not significant, so it is not discussed further. We first present analyses of prediction accuracy, followed by analyses of self-predictions and recall performance levels. Children's self- and other-predictions are then compared.

Prediction accuracy

Mean prediction accuracy across trials (predicted recall–actual recall) is presented in Table 1. Higher values indicate greater overconfidence. A 2 (Task: self-prediction accuracy or other-prediction accuracy) × 5 (Trial: 1, 2, 3, 4, or 5) × 2 (Order: self-prediction first or self-prediction second) × 2 (Video Model Gender: male or female) ANOVA yielded only a significant main effect of task, $F(1, 28) = 6.30$, $p < .05$, $\eta^2 = .18$, and no significant interactions. Other-predictive accuracy was greater (predictions were less overconfident) than self-predictive accuracy. Perhaps most important, in both conditions the children were substantially overconfident on the first trial and remained overconfident on the final trial. Thus, although the small effect of task (other-prediction vs. self-prediction) suggests that wishful thinking can account for some of the children's overconfidence, it cannot entirely explain their persistent overconfidence across all trials of this memory task.

Mean prediction and recall performance

Because prediction accuracy is derived from prediction and recall magnitudes, we also present these values in Table 2. Across all trials and participants, intrusion errors were minimal (only 16) and so were not included in these analyses. A 2 (Task: prediction or recall) × 5 (Trial: 1, 2, 3, 4, or 5) ANOVA yielded significant main effects of task, $F(1, 31) = 4.68$, $p < .05$, $\eta^2 = .72$, and trial, $F(4, 31) = 51.63$, $p < .01$, $\eta^2 = .16$, but no significant interaction. Consistent with the analysis of prediction accuracy, children's predicted recall (overall $M = 6.18$) was substantially higher than what they actually recalled (overall $M = 2.81$).

Both prediction and recall values declined slightly across the five trials. The mean prediction on the first trial was significantly different from that on the fifth trial, $t(31) = 2.53$, $p < .05$, $d = .48$. The mean prediction on the fourth trial was also significantly different from that on the fifth trial, $t(31) = 2.06$, $p < .05$, $d = .23$. No other differences in prediction were statistically significant, ts < 1. The mean recall on the first trial was significantly different from that on the second trial, $t(31) = 4.51$, $p < .01$, $d = .98$. No other differences in recall were statistically significant, ts < 1.

Comparison of self- and other-predictions

The mean prediction values for self-predictions and other-predictions are presented in Table 2. A 2 (Condition: self or other) × 5 (Trial: 1, 2, 3, 4, or 5) ANOVA yielded no significant effects. Both self- and other-predictions declined over the five trials, but no significant differences occurred between the self- and other-predictions on any of the five trials, ts < 2.0. Thus, in an extension of Schneider (1998), children's predictions across trials were equally high in both the self and other conditions.

Correlations between self-predictions and recall

Correlations between the self-prediction values and the recall values are presented in Table 4. Predictions on the fifth trial were significantly correlated with recall performance on the fourth trial, $r(29) = .41$, $p < .05$. This result was the only instance in which the level of prediction on one trial was related to the level of recall on the preceding trial and so may be just a chance occurrence. Alternatively, it may be evidence that children's recall performance on the fourth trial influenced their prediction on the final trial. In Experiment 1, no significant correlations between any of the prediction values and recall values were observed. Thus, the two additional trials in Experiment 2 may have helped children to incorporate task experience into their predictions. With even more task experience (additional trials), children's recall and predictions may continue to be correlated. Finally, as in the

Please cite this article in press as: Lipko, A. R., et al. Persistent overconfidence despite practice: The role ... *Journal of Experimental Child Psychology* (2008), doi:10.1016/j.jecp.2008.10.002

Table 4
Correlations between self-predictions and recall with age partialled out in Experiment 2

	P2	P3	P4	P5	R1	R2	R3	R4	R5
Prediction 1	.42*	.51**	.21	.39*	.38*	−.19	.01	.13	−.05
Prediction 2		.35	.40*	.43*	**.25**	−.14	−.15	.07	−.25
Prediction 3			.56**	.62**	.22	**−.13**	.14	.20	−.27
Prediction 4				.79**	.14	−.22	**.21**	.41*	−.01
Prediction 5					.16	−.12	.27	**.41***	−.09
Recall 1						.11	.45*	.19	.15
Recall 2							.44*	.29	.14
Recall 3								.33	.38*
Recall 4									.22

Note. P2, prediction on Trial 2; P3, prediction on Trial 3; P4, prediction on Trial 4; P5, prediction on Trial 5; R1, recall on Trial 1; R2, recall on Trial 2; R3, recall on Trial 3; R4, recall on Trial 4; R5, recall on Trial 5. Bold values are the correlations between recall on Trial *N* and predictions on Trial *N* + 1.

* $p < .05$.

** $p < .01$.

Experiment 1, the correlations between predictions and recall on a given trial (which is a between-participants correlational measure of judgment accuracy) were typically near zero.

Experiment 3

In the first two experiments, preschoolers overconfidently predicted their picture recall. This overconfidence persisted across multiple trials despite the children's repeated experience of recalling fewer pictures than they had predicted. As discussed above, wishful thinking cannot account entirely for children's overconfidence, so in Experiment 3 we evaluated alternative explanations for their persistent overconfidence. One alternative for children's overconfidence on a memory task involves a memory monitoring deficiency in which young children do not accurately monitor their performance on a task (Schneider, 1998). As a result, they continue to give predictions that are well above their demonstrated performance level. Schneider (1998) evaluated this hypothesis by asking 4- and 6-year-olds to report their performance on the final trial on a location memory task. These postdictions were quite accurate, suggesting that children can monitor their memory performance accurately (see also Bjorklund, Gaultney, & Green, 1993).

If children can accurately monitor their performance on a task, why does their overconfidence persist across trials? In a multitrial task, monitoring previous performance is only one part of making an accurate prediction. It is also essential to correctly remember this information when making the prediction and to use it to temper one's overly optimistic expectations. In Experiments 1 and 2 of the current investigation, children did not appear to incorporate outcomes from previous trials into their predictions, as indicated by the null correlations between recall performance on Trial $N - 1$ and predictions on Trial *N* (see Tables 3 and 4). One exception was on Trial 5 in Experiment 2, which may be evidence that task performance that is less than expected can help to lower children's expectations for future performance. Assuming that this adjustment occurred in Experiment 2, it was far from adequate given that children's predictions on the final trial were still substantially overconfident.

Even if children can remember how well they performed on the previous recall trial, they still might not spontaneously think about such information and/or incorporate it into their predictions of future performance (e.g., Kail, 1990). Pressley, Ross, Levin, and Ghatala (1984) asked 10- and 13-year-olds to learn vocabulary words using one of two strategies. Children were asked to select one strategy to use during the learning process. Although they were aware of the relative efficacy of each strategy, they did not select the more effective one. Thus, remembering relevant information from one's previous experience does not guarantee that one will use such information in subsequent judgments and decisions (Pressley & Ghatala, 1989; Stipek et al., 1984).

In Experiment 3, we evaluated (a) whether children remembered the number of pictures they had recalled on the previous trial immediately before making their prediction and, if they did remember this number, (b) whether they used this information to adjust their predictions. To address these questions, the procedure was changed so that immediately before making a prediction on a given trial, some children were asked to postdict their performance on the immediately preceding trial. The performance of this postdiction group was compared with the performance of a control group whose members did not make postdictions prior to their predictions. This design allowed us to investigate the two possibilities described above. First, when making a postdiction, if information about their previous recall performance is available from memory, children's postdictions should be accurate. Second, if their postdictions are accurate and children use this information to update their predictions, prediction accuracy should be better for the postdiction group than for the control group. Overconfidence for the former group may even be diminished entirely.

Method

Participants

A total of 64 children (mean age = 4 years 11 months, range = 4 years 0 months to 5 years 11 months, 29 girls and 35 boys) were recruited to participate. None of the children had participated in either of the previous experiments. They were recruited through their preschools, which were located in a primarily suburban middle-class section of northeastern Ohio.

Design, materials, and procedure

Children were tested individually in a quiet room in their preschool. Before beginning the task, they were asked to count to 10 and to count the number of magnets on the board in front of them. They were then asked a few practice questions to ensure that they could compare quantities of small groups of objects. At the end of the session, the children were given a small gift for their participation.

The materials were identical to those used in the Experiment 2 (see Appendix A for lists of pictures). The procedure was the same as the one used for the self-prediction task in Experiment 2 but with two important differences. First, children were randomly assigned to one of two groups: postdiction or control. Second, they were asked a question immediately before predicting their recall on Trials 2, 3, 4, and 5. This question differed according to group. In the postdiction group, the children were asked, "How many pictures did you remember last time?" To use the same amount of time prior to making their predictions, the control group was asked, "Do you remember what we did last time?" In both groups, immediately after the children answered this question, the experimenter said, "Now, let's talk about these pictures [pointing to the pictures on the board]. How many pictures do you think you will be able to remember once I cover them up?" Thus, on each trial, children studied 10 new pictures, answered a question (postdiction or control) about the previous trial, made a prediction for the current trial, and then attempted recall. Note that a slightly modified version of this procedure occurred on Trial 1 because there was not a previous trial to refer to at the time of the question (postdiction vs. control).

This procedure was chosen to maximize the salience of previous trial performance when the children made their predictions. We hypothesized that making a postdiction immediately prior to making a prediction for a subsequent trial would lower children's expectations for future performance. Moreover, given this procedure, if children do not adjust their predictions, it cannot be because they simply forgot their previous postdiction but instead would likely result from them not considering this aspect of past performance when making their predictions.

Results and discussion

All analyses were first conducted with gender as a factor, but the main effect of gender and all interactions including it were not significant. Absolute accuracy for predictions and postdictions are presented first, followed by analyses of mean performance levels of predictions, recall, and postdictions. Predictions in the two groups (postdiction and control) are then compared to address whether having children make postdictions prior to making predictions influenced their levels of overconfidence.

Judgment accuracy

To evaluate whether children could accurately remember how many pictures they had recalled on each trial, we first analyzed the absolute accuracy of the postdictions. Absolute accuracy was computed by subtracting recall performance on a given trial (N) from each child's postdiction made on the immediately following trial ($N + 1$). (Note that absolute accuracy of postdictions cannot be computed for Trial 1 because a recall trial did not occur before the first study trial.) Mean absolute accuracy across the trials was as follows (with standard errors in parentheses): Trial 2 = 1.00 (0.44), Trial 3 = 0.94 (0.43), Trial 4 = 0.69 (0.37), and Trial 5 = 0.79 (0.36). An ANOVA revealed no significant main effects or interactions, $F\text{s} < 1.0$. Note that 17 of 33 children made correct postdictions on every trial and that 6 others made correct postdictions on three of four trials. Only 5 children made a single postdiction that was lower than their actual recall on the previous trial. Thus, the level of accuracy here is excellent (especially on the final trial). Such high levels of accuracy indicate that immediately before the children made their predictions, they could correctly remember how well they had performed on the previous trial.

Absolute accuracy for the predictions (predicted recall–actual recall) is presented in Table 1. Higher values indicate greater overconfidence. Absolute accuracy did not change across trials. In a replication of Experiment 1, a 2 (Group: postdiction or control) × 5 (Trial: 1, 2, 3, 4, or 5) ANOVA did not reveal any reliable effects or interactions, $F\text{s} < 3.0$. Thus, children remained overconfident across all trials, even children in the postdiction group who had just accurately remembered how many pictures they had recalled before making their predictions.

Prediction, recall, and postdiction magnitudes

Because accuracy is derived from judgment and recall magnitudes, we also present these values in Table 2. Across all trials and participants, intrusion errors were minimal (only 4) and so were not included in these analyses. A 2 (Task: prediction or recall) × 5 (Trial: 1, 2, 3, 4, or 5) × 2 (Group: postdiction or control) mixed ANOVA of children's predictions and recall yielded significant main effects of task, $F(1, 61) = 176.01$, $p < .01$, $\eta^2 = .74$, and trial, $F(4, 61) = 27.47$, $p < .01$, $\eta^2 = .31$, but no significant interactions. As in the previous two experiments, children's predicted recall on a trial (overall $M = 7.61$) was substantially higher than what they actually recalled on that trial (overall $M = 3.56$).

Self-predicted and actual recall declined over trials. The mean prediction on the first trial was significantly greater than that on each of the four subsequent trials, $t\text{s} > 2.64$, $p < .05$, $d\text{s} > .64$. The mean prediction on the second trial was also significantly greater than that on the third trial, $t(62) = 2.27$, $p < .05$, $d = .24$. No other differences in prediction magnitudes were statistically significant, $t\text{s} < 1$. The mean recall on the first trial was significantly greater than that on each of the four subsequent trials, $t\text{s} > 6.72$, $p < .01$, $d\text{s} > .92$. The mean recall on the second trial was also significantly greater than that on the fourth trial, $t(63) = 2.62$, $p < .01$, $d = .39$. No other differences in recall were statistically significant, $t\text{s} < 1$.

Contrary to expectations, the effect of group was not significant, $F(1, 61) < 1$. Moreover, for the postdiction group, Trial N postdictions were significantly less than Trial $N + 1$ predictions on each of the four trials: Trial 1, $t(29) = 3.64$, $p < .05$, $d = .73$; Trial 2, $t(32) = 4.24$, $p < .001$, $d = .92$; Trial 3, $t(31) = 5.58$, $p < .001$, $d = 1.36$; Trial 4, $t(32) = 4.93$, $p < .001$, $d = 1.28$. Moreover, Trial N postdictions were not significantly correlated with $N + 1$ predictions on any trial ($r\text{s} = -.05, .27, .06$, and $.04$, respectively). Thus, even though children's postdictions tended to be accurate, their predictions were not significantly related to the predictions they provided for the next trial. This evidence indicates that children did not use information from postdictions to adjust their subsequent predictions.

Correlations between predictions and recall

Correlations between self-prediction values and recall values with age partialled out are presented in Table 5. The number of pictures that children predicted they would recall on a particular trial was not correlated with the number they actually recalled on the previous trial, with one exception: A small significant positive correlation occurred between recall on Trial 3 and prediction for Trial 4 ($r = .27$). Although in Experiment 2 recall on the fourth trial was significantly correlated with predictions on the fifth trial, this finding was not replicated in the current experiment. Finally, as in the pre-

Please cite this article in press as: Lipko, A. R., et al. Persistent overconfidence despite practice: The role ... *Journal of Experimental Child Psychology* (2008), doi:10.1016/j.jecp.2008.10.002

Table 5
Correlations between self-predictions and recall (collapsed across condition) with age partialled out in Experiment 3

	P2	P3	P4	P5	R1	R2	R3	R4	R5
Prediction 1	.35**	.25	.35**	.43**	−.04	−.09	.01	.04	−.20
Prediction 2		.59**	.50**	.47**	**−.01**	.01	.17	−.08	−.20
Prediction 3			.72**	.68**	.08	**.16**	.26*	−.05	.00
Prediction 4				.85**	.08	.22	**.27***	−.04	−.06
Prediction 5					.14	.16	.24	**.07**	−.05
Recall 1						.31*	.36**	.45**	.23
Recall 2							.33**	.26*	.29*
Recall 3								.35**	.17
Recall 4									.25

Note. P2, prediction on Trial 2; P3, prediction on Trial 3; P4, prediction on Trial 4; P5, prediction on Trial 5; R1, recall on Trial 1; R2, recall on Trial 2; R3, recall on Trial 3; R4, recall on Trial 4; R5, recall on Trial 5. Bold values are the correlations between recall on Trial *N* and predictions on Trial *N* + 1.
* $p < .05$.
** $p < .01$.

vious experiments, the correlations between predictions and recall on a given trial were typically near zero and not statistically significant.

General discussion

Across three experiments, children's poststudy recall predictions were persistently overconfident despite repeated experience with the recall task. Given that the sample sizes in the first two experiments were relatively small, potentially subtle changes in overconfidence across trials might not have been detected. To evaluate this possibility, we examined prediction accuracy for all of the control groups (whose members made self-predictions only) collapsed across experiments (combined $N = 84$). Mean prediction accuracy across the first three trials (which were included in all experiments) were as follows (with standard errors in parentheses): 3.18 (0.29), 4.14 (0.33), and 3.70 (0.33) on Trials 1, 2, and 3, respectively, $F(2, 166) = 3.94$, $p < .05$, $\eta^2 = .05$. Post hoc analyses revealed that predictive accuracy on Trial 1 was actually worse than that on Trial 2, $t(83) = -2.89$, $p < .01$, $\eta^2 = .34$, but not significantly different from predictive accuracy on Trial 3 ($t < 2$). Children remained overconfident across all three trials. In the remainder of this section, we consider reasons why children may remain persistently overconfident across multiple trials of the same recall task, and then we discuss the importance of overconfidence to adaptive behavior.

Explaining children's persistent overconfidence despite practice

Wishful thinking?

In three experiments, preschoolers overconfidently predicted their picture recall. Previously, the wishful thinking hypothesis (Schneider, 1998; Stipek et al., 1984) has been offered as an explanation for such overconfidence. According to this hypothesis, children base their predictions on their desire for performance rather than on their actual expectations. The hypothesis assumes that children desire higher levels of performance for themselves and that, as a result, their overconfidence is greater when predicting their own performance than when predicting another child's performance.

Results of Experiment 2 suggest that this hypothesis may partially account for children's overconfidence on the current task. As shown in Table 1, overconfidence was somewhat greater for self-predictions than for other-predictions. Nevertheless, both self- and other-predictions were overconfident across all trials (Table 1), and the magnitudes of self- and other-predictions did not differ significantly (Table 2). Both of these outcomes are inconsistent with the wishful thinking hypothesis, suggesting that it cannot fully account for children's persistent overconfidence across trials. These findings extend the results of Schneider (1998), who used a substantively different memory task. Thus, although wishful thinking may be a viable explanation for children's overconfidence on psychomotor tasks (Schneider, 1998; Stipek et al., 1984), it is not the sole explanation for their overconfidence on the current memory task.

Given all of the potential differences between psychomotor and cognitive tasks, there are many reasons why wishful thinking may contribute differentially to overconfidence on the two task types. As compared with cognitive tasks, children may be more likely to view psychomotor tasks as a game and become more competitive; hence, they may want others to perform worse. Alternatively, the psychomotor tasks may be consistently easier (e.g., most children can throw a ball or jump); hence, children may believe that they should be able to perform the tasks well. Of course, these explanations are speculative and, along with other explanations, will require systematic research to evaluate.

Memory for previous performance and using this information in updating predictions

An alternative explanation for children's persistent overconfidence was evaluated in Experiment 3. Specifically, some children were asked to make a postdiction about their recall on the previous trial before making a prediction for the current trial. Consistent with previous research (Schneider, 1998), children made accurate postdictions, suggesting that they could accurately remember how they had performed on the previous trial when making a new prediction. Nevertheless, children's predictions in the two groups (postdiction and control) did not differ, and the children who made postdictions prior to making their predictions were not less overconfident across trials. Thus, even though making postdictions may have caused children's own recall performance to be more salient when making predictions, such changes in salience did not influence their predictions.

Why might children neglect past performance when making predictions? Young children do not commonly evaluate their own abilities and, thus, do not have much experience in adjusting their expectations based on previous experience. In early education settings, teachers tend to focus on learning goals rather than on performance goals. Mastery of skill is encouraged rather than competition among children (e.g., Midgley, Kaplan, & Middleton, 2001). Thus, children may be rewarded for trying a task and persisting at it regardless of how well they actually perform it. Although such reinforcement patterns are likely to promote motivation and self-esteem, they would not give children experience with true evaluations of their abilities.

Moreover, using previous experience to predict future performance may be a learned strategy, and young children tend to be less likely to spontaneously generate and use beneficial strategies compared with older children (e.g., Kail, 1990; Kreutzer, Leonard, & Flavell, 1975). In addition to failing to appreciate the benefits of strategies, young children often balk at expending the effort to use a strategy. It can be difficult to learn to use a strategy correctly, and in certain situations the costs may outweigh the benefits (Siegler & Alibali, 2005). Overconfidence may further encourage this tendency because if children are unaware that they need a strategy, they will be even less likely to search for one or implement one. Finally, the ability to monitor performance accurately and the ability to use previous performance to predict future performance may develop at different times. Translating past performance into future performance may require more cognitive resources and, hence, may take more time to develop. These explanations should be evaluated in future research, but most important, the current experiments have isolated children's persistent overconfidence to their failure to use accurate memory for past performance to adjust their subsequent predictions.

Is children's overconfidence adaptive?

Although overconfidence declines during childhood (Schneider & Pressley, 1997), older children and adults are also overly optimistic about their abilities in certain situations (Metcalfe, 1998; Plumert, 1995; Plumert & Schwebel, 1997). For example, middle school and college students have been found to be significantly overconfident when asked to evaluate their learning of text materials even when provided with feedback (Lipko et al., 2007). This type of overconfidence can lead to negative learning outcomes. Students will cease studying materials that they believe they have already learned successfully, so if they stop studying prematurely due to inappropriate confidence, they may perform poorly on tests of the materials.

With respect to physical activities, overconfidence in the face of contradictory task experience can have harmful effects. For example, Plumert (1995) asked 6-year-olds, 8-year-olds, and adults to predict whether they would be able to successfully perform physical tasks that were either within or beyond the participants' actual ability levels. The children who made the most errors in judging their ability to accomplish physical tasks were the same children who had experienced more unintentional injuries compared with the children who had made fewer errors in judgment. Remaining overconfident in the face of feedback can also lead to persistent failure on some tasks, especially if young children believe that the amount of effort they exert on tasks is directly predictive of their success (Stipek & MacIver, 1989). Believing that effort, rather than their previous performance, is a good indicator of future performance, children are likely to remain overconfident and might not seek assistance to improve their performance.

Nevertheless, in some circumstances, children's overconfidence may be adaptive (Bjorklund et al., 1993; Bjorklund & Green, 1992; Shin et al., 2007). In fact, this overconfidence may be especially adaptive for younger children (e.g., Bjorklund et al., 1993). Overconfidence can help to maintain and protect children's self-esteem and may serve to encourage children to try new tasks on which they may initially perform poorly. Having an optimistic view of their abilities can encourage children to try to achieve loftier goals for themselves (Bandura, 1989). If children are motivated to persist at a task, they will subsequently gain practice, which itself may lead to increased performance (Shin et al., 2007).

Summary

In three experiments, young children's poststudy predictions of their recall performance were persistently overconfident. This overconfidence was maintained despite a great deal of task experience in which children performed substantially lower than they had predicted. Wishful thinking may provide a partial explanation for children's overconfidence, which appears to arise because children fail to use accurate memory for past performance in adjusting their subsequent predictions. Discovering why children fail to make such adjustments is an important goal for future research.

Acknowledgment

We thank the families and staff of the local preschools for all of their help.

Appendix A

A.1. Experiment 1 stimuli

Trial 1: blocks, bike, plant, horse, chair, lion, teddy bear, car, ball, sneakers
Trial 2: keys, shirt, shovel, bird, ice cream, cat, truck, doll, butterfly, shark
Trial 3: bucket, snake, toothbrush, sunglasses, coat, boat, dog, grapes, clock, cup.

A.2. Experiments 2 and 3 stimuli

Trial 1: cat, juice, car, sneakers, cookies, keys, teddy bear, crayons, bird, ice cream
Trial 2: butterfly, shirt, banana, heart, toothbrush, ball, snake, light bulb, shark, pig
Trial 3: gloves, carriage, dress, truck, balloon, shovel, doll, bucket, horse, sunglasses
Trial 4: lion, blocks, grapes, boat, rocks, train, cup, clock, bug, cherries
Trial 5: dog, paper, bike, chair, duck, house, fish, plant, coat, leaves.

References

Bandura, A. (1989). Regulation of cognitive processes through perceived self-efficacy. *Developmental Psychology, 25*, 729–735.
Bjorklund, D. F., Gaultney, J. F., & Green, B. L. (1993). I watch, therefore I can do: The development of meta-imitation during the preschool years and the advantage of optimism about one's imitative skills. In R. Pasnak & M. L. Howe (Eds.). *Emerging themes in cognitive development* (Vol. 2, pp. 79–102). New York: Springer-Verlag.
Bjorklund, D. F., & Green, B. L. (1992). The adaptive nature of cognitive immaturity. *American Psychologist, 47*, 46–54.
Cunningham, J. G., & Weaver, S. L. (1989). Young children's knowledge of their memory span: Effects of task and experience. *Journal of Experimental Child Psychology, 48*, 32–44.

Dunlosky, J., & Hertzog, C. (2000). Updating knowledge about encoding strategies: A componential analysis of learning about strategy effectiveness from task experience. *Psychology and Aging, 15*, 462–474.

Flavell, J. H., Friedrichs, A. G., & Hoyt, J. D. (1970). Developmental changes in memorization processes. *Cognitive Psychology, 1*, 324–340.

Hertzog, C., Dixon, R. A., & Hultsch, D. F. (1990). Relationships between metamemory, memory predictions, and memory task performance in adults. *Psychology and Aging, 5*, 215–227.

Hertzog, C., Saylor, L. L., Fleece, A. M., & Dixon, R. A. (1994). Metamemory and aging: Relations between predicted, actual, and perceived memory task performance. *Aging and Cognition, 1*, 203–237.

Kail, R. (1990). *The development of memory in children* (3rd ed.). New York: W. H. Freeman.

Kreutzer, M. A., Leonard, C., & Flavell, J. H. (1975). An interview study of children's knowledge about memory. *Monographs of the Society for Research in Child Development, 40*, 1–60.

Lipko, A. R., Dunlosky, J., Rawson, K. A., Swan, K., & Cook, D. (2007). *Using feedback to improve undergraduate and middle-school students' metacomprehension accuracy*. Kent State University: Unpublished manuscript.

Metcalfe, J. (1998). Cognitive optimism: Self-deception or memory-based processing heuristics. *Personality and Social Psychology Review, 2*, 100–110.

Midgley, C., Kaplan, A., & Middleton, M. (2001). Performance-approach goals: Good for what, for whom, under what circumstances, and at what cost? *Journal of Educational Psychology, 93*, 77–86.

Plumert, J. M. (1995). Relations between children's overestimation of their physical abilities and accident proneness. *Developmental Psychology, 31*, 866–876.

Plumert, J. M., & Schwebel, D. C. (1997). Social and temperamental influences on children's overestimation of their physical abilities: Links to accidental injuries. *Journal of Experimental Child Psychology, 67*, 317–337.

Pressley, M., & Ghatala, E. S. (1989). Metacognitive benefits of taking a test for children and young adolescents. *Journal of Experimental Child Psychology, 47*, 430–450.

Pressley, M., Ross, K. A., Levin, J. R., & Ghatala, E. S. (1984). The role of strategy utility knowledge in children's strategy decision making. *Journal of Experimental Child Psychology, 38*, 491–504.

Schneider, W. (1998). Performance prediction in young children: Effects of skill, metacognition, and wishful thinking. *Developmental Science, 1*, 291–297.

Schneider, W., & Lockl, K. (2008). Procedural metacognition in children: Evidence for developmental trends. In J. Dunlosky & R. A. Bjork (Eds.), *Handbook of metamemory and memory* (pp. 391–409). New York: Taylor & Francis.

Schneider, W., & Pressley, M. (1997). *Memory development between two and twenty* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum.

Schneider, W., Visé, M., Lockl, K., & Nelson, T. O. (2000). Developmental trends in children's memory monitoring: Evidence from a judgment-of-learning (JOL) task. *Cognitive Development, 15*, 115–134.

Shin, H., Bjorklund, D. F., & Beck, E. F. (2007). The adaptive nature of children's overestimation in a strategic memory task. *Cognitive Development, 22*, 197–212.

Siegler, R. S., & Alibali, M. W. (2005). *Children's thinking* (4th ed.). Upper Saddle River, NJ: Prentice Hall.

Stipek, D. J., & Hoffman, J. M. (1980). Development of children's performance-related judgments. *Child Development, 51*, 912–914.

Stipek, D. J., & MacIver, D. (1989). Developmental change in children's assessment of intellectual competence. *Child Development, 60*, 521–538.

Stipek, D. J., Roberts, T. A., & Sanborn, M. E. (1984). Preschool-age children's performance expectations for themselves and another child as a function of the incentive value of success and the salience of past performance. *Child Development, 55*, 1983–1989.

Thiede, K. W. (1999). The importance of monitoring and self-regulation during multitrial learning. *Psychonomic Bulletin & Review, 6*, 662–667.

Wilkes, A. (1999). *My first word book*. New York: DK Publishing.

Yussen, S. R., & Levy, V. M. (1975). Developmental changes in predicting one's own span of short-term memory. *Journal of Experimental Child Psychology, 19*, 502–508.

Please cite this article in press as: Lipko, A. R., et al. Persistent overconfidence despite practice: The role ... *Journal of Experimental Child Psychology* (2008), doi:10.1016/j.jecp.2008.10.002