International Education Studies; Vol. 9, No. 3; 2016
ISSN 1913-9020 E-ISSN 1913-9039
Published by Canadian Center of Science and Education

Prediction Accuracy: The Role of Feedback in 6th Graders' Recall Predictions

Ibrahim S. Al-Harthy[1]

[1] Psychology Department, Sultan Qaboos University, Muscat, Oman

Correspondence: Ibrahim S. Al-Harthy, B.O.Box, 32, PC 123, Al-Khud, Muscat, Oman. Tel: 1-968-2414-3945. E-mail: ibrahimh@squ.edu.om

Received: August 30, 2015 Accepted: October 10, 2015 Online Published: February 25, 2016

doi:10.5539/ies.v9n3p212 URL: http://dx.doi.org/10.5539/ies.v9n3p212

Abstract

The current study focused on the role of feedback on students' prediction accuracy (calibration). This phenomenon has been widely studied, but questions remain about how best to improve it. In the current investigation, fifty-seven students from sixth grade were randomly assigned to control and experimental groups. Thirty pictures were chosen from the curriculum and these were the same in the trials for both groups, 10 for each trial. The pilot study showed all pictures had the same difficulty level and that they were appropriate for 6th graders. During three trials, students in the experimental group received feedback about their prediction and recall. The purpose of the feedback was to assist students in monitoring their learning. Results demonstrated that students in the experimental group showed improvement in predication and prediction accuracy in trials 2 and 3. Significant differences were found among groups in trials 2 and 3, but not in trial 1. Implications and future studies are recommended.

Keywords: metacognition, prediction, recall, prediction accuracy

1. Introduction

A considerable number of researchers have investigated metamemory (e.g., Al-Balushi & Al-Harthy, 2015; Doulosky & Metcalfe, 2009; Lipko, Dunlosky, Lipowski, & Merriman, 2011; Lipko, Dunlosky, & Merriman, 2009; Miller & Geraci, 2011b). Metamemory has been defined as one's awareness of memory abilities and has different components: 1) knowledge about one's own memory, 2) monitoring and control of learning. Given the relationship between memory monitoring and learning and subsequent retention (Dunlosky & Metcalfe, 2009), research has extensively explored this area. One question that has received high attention is how accurate are children' monitoring predictions? And can this monitoring accuracy be developed in classroom? Absolute accuracy (the match between the magnitude of predictions and the level of memory performance) has been widely used in research. For instance, young children (e.g., 4- to 7-year-olds) tend to be more overconfident in memory abilities than older children. The current study contributes to this broad literature by exploring children' absolute accuracy (in terms of prediction accuracy) when they have multiple trials of receiving feedback about their previous performance.

When children are asked to make self-assessments of their learning, they are usually overconfident. For example, children in grade 1-3 overestimated the number of pictures they can recall after studying them (Lipko et al., 2009, Lipko et al., 2011). Although prediction declines with age, research has found undergraduate students are also overconfident in their academic abilities. They tend to overestimate their performance on upcoming exams (Hacker, Bol, Horgan, & Rakow, 2000; Millar & Geraci, 2010). In Hacker et al. (2000) undergraduate students were asked to record their prediction of what percentage of an exam's items they would get correct before the exam. They found that students with higher scores had more accurate predictions than the students with lower scores.

Most studies have confirmed that students tend to overestimate their own ability (e.g., Al-Harthy, Was, & Hassan, 2015; Glenberg & Epstein, 1985; Glenberg, Sanocki, Epstein, & Morris, 1987; Lichtenstein, Fischhoff, & Philips, 1982). Children (5-7 year old) do not show any underconfidence even when they receive practice. The underconfidence pattern starts to arise in 9 years old children (Lipko et al., 2011). Children' prediction accuracy (PA) becomes much better when they perform tasks (3 trials) with same pictures, but overconfidence remains for new pictures (Lipko-Speed, 2013).

Memory-related variables might add to our understanding of children' overestimation in prediction tasks. Kvavilashvili and Ford (2014) listed these memory-related variables, which are task, person, and strategy. More specifically, the task variable is related to long vs. short memory list, in that long lists are more difficult to remember. The second variable is person, which focuses on age differences. This means that adults outperform children on memory tasks. Lastly, strategy is basically the method used during the learning process of the tasks, for example, rehearsing vs. looking. Children' limited understanding of these variables has played a role in their prediction (Bjorklund, Dukes, & Brown, 2009; Wellman, 1977).

Other studies have found that PA is often resistant to improvement (e.g., Gigerenzer, Hoffrage, & Kleinbolting, 1991; Koriat, 1997; Koriat, Lichtenstein, & Fischhoff, 1980; Lipko et al., 2009; Miller & Geraci, 2011a; Nietfeld, Cao, & Osborne, 2005). Results invariably show that 4 to 6 year-old children grossly overestimate the number of items they can recall. For example, Lipko et al. (2009) used study-predict-recall methodology (explained below) with 21 children. Across three trials, children remained overconfident although they were informed the number of pictures that they had successfully remembered in previous trials. This demonstrated that preschoolers' overconfidence was remarkably resistant to the repeated experience of recalling fewer pictures.

1.1 Rational for the Current Study

The current study argues that metamemory abilities can be learned and improved. According to Baker and Brown (1984), two components are involved in metacognition: 1) students' awareness of the task to be completed, and 2) students' cognitive monitoring. In the current study, students in the experimental group received feedback about their performance in previous trials. This provided the experimental group with a chance to monitor, review, and determine if the task was being completed properly and to make corrections as appropriate. Also, additional information about upcoming trials (2 & 3) was provided before participants started trials 2 and 3. This offered some information about the task to be completed.

Although the current study used the same experimental paradigms (study-predict-recall) used in Lipko's et al. (2009) and Lipko-Speed (2013), age differences exist between the current experiment and Lipko's. In addition, in trials 2 and 3 in the current study, participants were told the number of pictures they had correctly recalled. Also, this number was visually presented to participants on the participant's experiment sheet. Moreover, they were informed that the pictures they would be studying and trying to recall (in trials 2 and 3) differed among the trials but were at the same difficulty level as those in trial 1.This was an attempt to control for the children' unawareness of pictures to be recalled in upcoming trials. In the current study, trial one's goal was to replicate previous findings in that both groups (control and experimental) would be overconfident in their PA. In trials 2 and 3, the goal was to evaluate whether children' overconfidence in the experimental group would differ with treatment manipulation. A gap that has been identified in this research is whether 6th graders' overconfidence in memory performance would remain when feedback was presented about the memory task stimuli? In short, would participants' PA in the experimental group continuously get better when feedback was provided?

Importantly, exploring children' prediction across trials allows for a better evaluation for this phenomenon. For example, the wishful thinking hypothesis suggests that a child's prediction is based on how s/he wants to perform rather than on how s/he expects to perform (Stipek, Roberts, & Sanborn, 1984). If correct, this would tell us that children' overconfidence would remain equal across all three trials in both groups.

2. Method

2.1 Participants

57 participants from grade six (age between 11-12 years old, meals = 27 and females = 30) were recruited from a single public elementary school in Muscat, Oman. The recruitment procedure was that grade 6 teachers informed students about the research and volunteers were given gifts for their participation. Three participants were eliminated from the analysis because their prediction accuracy measure could not be computed. These subjects always used the maximum confidence rating. All participants had equivalent instructions and had the same grade requirements.

2.2 Materials

The material for this study consisted of 30 pictures chosen from the grade six curriculums in the Sultanate of Oman. Pictures were reviewed by school teachers to ensure equivalence with regards to difficulty. Also, pilot testing determined that the pictures were familiar to the 6th graders and were at the same difficulty level.

2.3 Procedure/Design

Typically, children' limited knowledge has been assessed by the study-predict-recall methodology (Kvavilashvili

& Ford, 2014; Lipko et al., 2012). In this method, children are exposed to the to-be-recalled materials and are asked to predict how many they will be able to recall from memory before actually recalling them. In the current study, participants were randomly assigned to either a control or an experimental group. Both groups received three trials. Participants from both groups were tested individually in a quiet room in the school. In trial 1 and after the experiment instructions had been given, each participant was seated in front of a large white board and told that 10 pictures would be presented one at a time. The 10 pictures were the same for both groups in the same trial, but differed between trials. The subject was instructed to pay attention to the picture, name it, and try to remember as many as possible. The experimenter placed one picture at a time and asked the participant to name the picture as it was presented. The experimenter continued presenting pictures with already named pictures left uncovered. After presenting all 10 pictures and placing them on the board, additional time (15s) was given for the participant to study the 10 pictures. Then, the participant was asked to predict how many pictures s/he could recall when the pictures were covered. After the study phase and the prediction judgment were made, the experimenter covered the pictures and asked the participant to recall them. The number of correct pictures recalled out of 10 was recorded in the participant's experiment sheet. After 1 minute elapsed, the second trial began: participants in the experiment group received feedback about their first trial result. They were informed about their prediction and the number of correct pictures recalled. For example, a subject was told "*you predicted that you would remember 8 pictures and you successfully recalled 6 pictures*". Also, participants in the experiment group were told that all pictures in 3 trials had same difficulty level. The purpose of offering feedback was to help the participant reflect and analyze why their predictions were or were not accurate. Participants in the control group did not receive this feedback. The procedure in trials 2 and 3 were similar to those in trial 1 with feedback about previous trial offered to participants in the experimental group but not the control group.

3. Results

Difference scores for each student were created by subtracting the recall from the predicted number of pictures. As a result, positive numbers indicated overconfidence and negative numbers indicated underconfidence (see table 1). As expected in trial 2 and 3, the omnibus F test indicated significant differences in prediction accuracy by groups, $F(1,56) = 4.94$, $p < .05$; $F(1,56) = 4.15$, $p < .05$, respectively.

Participants in the experimental group were more accurate in trials 2 and 3 compared to those of the control group. In addition, the omnibus F test indicated significant difference in prediction judgment by groups. In trial 2, $F(1,56) = 8.93$, $p < .05$. In trial 3, $F(1,56) = 7.27$, $p < .05$. Participants in the experimental group lowered their prediction in trials 2 and 3 compared to subjects in the control group. Importantly, no significant differences in prediction or accuracy were found between groups in trial 1.

Table 1. Means and standard deviations for prediction, recall, prediction accuracy among groups

	Control Group			Experimental Group		
Trials	P	Recall	PA	P	Recall	PA
1	9.22 (1.1)	8.74 (1.3)	.48 (1.2)	8.67 (1.1)	8.57 (1.1)	.10 (1.3)
2	9.07 (1.4)	7.78 (1.0)	1.29 (1.5)	7.90 (1.5)	7.73 (1.4)	.16 (2.1)
3	9.41 (1.1)	7.96 (1.1)	1.4 (1.2)	8.43 (1.5)	7.80 (1.2)	.63 (1.7)

P = Prediction, PA= Prediction Accuracy. Numbers in parenthesis = SD.

4. Discussion

In contrast to studies that have shown no improvement in children' PA, participants in the experimental group in the present study demonstrated better performance in PA compared to the control group. Although no significant differences in PA was found between groups in trial 1, the PA pattern changed in trials 2 and 3. More specifically, participants' PA (calibration) in the experimental group was .16 and .63 in trials 2 and 3 respectively, whereas PA for the control group was 1.29 and 1.4 respectively. Also, participants' prediction in trial 1 did not significantly differ among groups. After treatment, participants' prediction in trial 2 and 3 differed significantly among groups. Participants in the experimental group lowered their prediction in trial 2 and 3, whereas those in the control group continued to overestimate their prediction ability.

Students' judgments (prediction) concerning future performance are influenced by (a) their understanding of how

much learning has occurred (b) how much learning might be forgotten (c) how difficult the task is. In the current study, feedback was provided to the experimental group. The feedback presented information regarding participant's performance. This performance included prediction and recall. This information assisted the participants to focus on two questions: what is learnt? And has learning occurred? Also, this information gave feedback on how much task was being forgotten or how much could not be recalled. In short, the feedback assisted the participant through the prediction process. On the other hand, participants in the control group continued to be overconfident in their prediction. This can be explained by the wishful thinking hypotheses. Their prediction was based on their desire to perform, not on their expectation. They exhibited poor self-assessment of how much learning occurred and continuously overestimated their prediction ability. The fact that participants in the control group were less accurate in their prediction than were participants in the experimental group supports the notion that metamemory ability can be improved, thus students can monitor their learning. Interestingly, this study contributes to the literature in that 6th graders might not be able to monitor their learning by themselves, but they demonstrate improvement in their learning monitoring when assistance is provided. One implication that the current study sheds lights on is that teachers play a main role in classroom settings. Teacher instructions are highly important, not only to finish classroom tasks, but also to assist students to monitor their learning.

Johnson-Laird (1983) argued that one's ability to focus on what is important necessarily requires the construction of mental models. A mental model combines knowledge and information about the current situation, thus it serves two purposes: it presents abstract relations between items in the situation and importantly it directs attention toward information that will further develop the mental model. In the present study, participants in the control group fail to form a mental model of experiment trials. Most likely, 6th graders fail to attend to what the trials are presenting to form the relationship between predictions and recall among trials, which indeed causes failure in constructing a situation model of the trials. On the other hand, participants in the experimental group were assisted by feedback that directed their attention to building a mental model of the situation. As presented above, this helped participants in the experimental group to keep their attention on what was relevant and important to monitor learning.

Given the current results showing differences in prediction accuracy among groups due to feedback, future research might also consider the contribution of teachers' feedback to students' performance. For example, what type of feedback is needed in classroom? And how do different types of feedback work in relation to the monitoring of learning?

Acknowledgements

This research was fully supported by the Sultan Qaboos University.

References

Al-Balushi, S. M., & Al-Harthy, I. S. (2015). Students' mind wandering in macroscopic and submicroscopic textual narrations and its relationship with their reading comprehension, *Chemistry Education Research and Practice, 16*, 680-688. http://dx.doi.org/10.1039/C5RP00052A

Al-Harthy, I. S., Was, C. A., & Hassan, A. S. (2015). Poor performers are poor predictors of performance and they know it: Can they improve their prediction accuracy? *Journal of Global Research in Education and Social Science, 4*(2). 93-100

Baker, L., & Brown, A. L. (1984). Metacognitive Skills and Reading. In P. D. Pearson, R. Barr, M. L. Kamil, & P. Mosenthal (Eds.), *Handbook of Reading Research* (pp. 353-394). New York: Longman.

Bjorklund, D. F., Dukes, C., & Brown, R. D. (2009). The development of memory strategies. In M. L. Courage, & N. Cowan (Eds.), *The development of memory in infancy and childhood* (pp. 146-175). Hove, UK: Psychology Press.

Dunlosky, J., & Metcalfe, J. (2009). *Metacognition*. Thousand Oaks, CA: Sage Publications.

Gigerenzer, G., Hoffrage, U., & Kleinbolting, H. (1991). Probabilistic mental models: A Brunswikian theory of confidence. *Psychological Bulletin, 98*, 509-528. http://dx.doi.org/10.1037/0033-295x.98.4.506

Glenberg, A. M., & Epstein, W. (1985). Calibration of comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 11*, 702-718. http://dx.doi.org/10.1037/0278-7393.11.1-4.702

Glenberg, A. M., Sanocki, T., Epstein, W., & Morris, C. (1987). Enhancing calibration of comprehension. *Journal of Experimental Psychology: General, 116*, 119-136. http://dx.doi.org/10.1037/0096-3445.116.2.119

Hacker, D. J., Bol, L., Horgan, D. D., & Rakow, E. A. (2000). Test Prediction and Performance in a Classroom Context. *Journal of Educational Psychology, 92*, 160- 170. http://dx.doi.org/10.1037/0022-0663.92.1.160

Johnson-Laird, P. N. (1983). Mental models: Towards a cognitive science of language, inference, and consciousness. *Journal of Experimental Psychology: General, 111*, 228-238.

Koriat, A. (1997). Monitoring one's won knowledge during study: A cue-utilization approach to judgments of learning. *Journal of Experimental Psychology: General, 126*, 349-370. http://dx.doi.org/10.1037/0096-3445.126.4.349

Koriat, A., Lichtenstein, S., & Fischhoff, B. (1980). Reasons for confidence. *Journal of Experimental Psychology: Human Learning and Memory, 6*, 107-118. http://dx.doi.org/10.1037/0278-7393.6.2.107

Kvavilashvili, L., & Ford, R. M. (2014). Metamemory prediction accuracy for simple prospective and retrospective memory tasks in 5-year-old children. *Journal of Experimental Child Psychology, 127*, 65-81. http://dx.doi.org/10.1016/j.jecp.2014.01.014

Lichtenstein, S., Fischhoff, B., & Phillips, L. D. (1982). Calibration of probabilities: The state of the art to 1980. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 306-334). New York: Cambridge University Press.

Lipko, A. R., Dunlosky, J., & Merriman, W. E. (2009). Persistent overconfidence despite practice: The role of task experience in preschoolers' recall predictions. *Journal of Experimental Child Psychology, 103*, 152-166. http://dx.doi.org/10.1016/j.jecp.2008.10.002

Lipko, A. R., Dunlosky, J., Lipowski, S. L., & Merriman, W. E. (2011). Young children are not underconfident with practice: the benefit of ignoring a fallible memory heuristic. *Journal of Cognition and Development, 13,* 174-188. http://dx.doi.org/10.1080/15248372.2011.577760

Lipko-Speed, A. R. (2013). Can young children be more accurate predictors of their recall performance? *Journal of Experimental Child Psychology, 114*, 357-363. http://dx.doi.org/10.1016/j.jecp.2012.09.012

Miller, T. M., & Geraci, L. (2011a). Unskilled but aware: reinterpreting overconfidence in low-performing students. *Journal of Experimental Psychology: Learning Memory, and Cognition, 37*(2), 502-506. http://dx.doi: 10.1037/a0021802

Miller, T. M., & Geraci, L. (2011b). Training Metacognition in the Classroom: The Influence of Incentives and Feedback on Exam Predictions. *Metacognition and Learning, 6*, 303-314. http://dx.doi.org/10.1007/s11409-011-9083-7

Nietfeld, J. L., Cao, L., & Osborne, J. W. (2005). Metacognitive Monitoring Accuracy and Student Performance in a Postsecondary Classroom. *The Journal of Experimental Education, 74*, 7-28.

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. http://dx.doi.org/10.2307/1129773

Wellman, H. M. (1977). Preschoolers' understanding of memory relevant variables. *Child Development, 48*, 1720-1723. http://dx.doi.org/10.2307/1128544