Towards a Developmental Understanding of Promisingness Judgments [1]

Bodong Chen, Marlene Scardamalia, Carl Bereiter

OISE/University of Toronto

[1] Manuscript prepared for 2013 AERA Annual Meeting, San Francisco, April 27 - May 1, 2013. Correspondence should be addressed by email to bodong.chen@utoronto.ca.

Abstract: This study examines the developmental aspect of promisingness evaluation. To this end, we designed a survey and invited participants to evaluate a list of ideas according to a multilevel promisingness scale. Students from Grade 4, 5 and 7, together with adult non-experts and domain experts, participated in this study. Results indicated domain knowledge played an important role in promisingness judgments; nevertheless, young students could perform fairly well as long as they acquire some basic content knowledge. Correlation analysis found the capability of promisingness judgments significantly correlated with age and domain knowledge; however, it was unclear which factor was more dominant in promisingness judgments. Future research with revised instruments is needed to address this issue.

Introduction

The capability of recognizing promisingness is one key manifestation of "creative expertise"[2]. In Dunbar's study of scientfic reasoning in real-world laboratories, he finds that scientists tend to evaluate their projects in terms of risk and promisingness; they are keen to work on promising projects that had the prospect of being an important discovery although it may also have high probability of failure[3]. Similarly, in explaining creative processes, Howard Gardner raises attention to the step of using intuition to detect "anomalty" or "discrepancy" when working creatively within a domain; "promisingness", as he explains, is what makes those discrepant elements stand out and further encourages individuals to "cast around" for ways to deal with them and to ultimately achieve local-coherence[4]. The same story could be told of experts in various fields such as chess playing[5], engineering design[6], competitive sports[7], and so forth. Notwithstanding its importance, however, the term "promisingness" has seldom appeared beyond the knowledge building community and there is little research to address this concept.

Creative experts rely on their "knowledge of promisingness"—knowledge acquired over time as they engage in creative practices—to identify promising directions[8]. Thus, to help young students develop such knowledge, they also need to work creatively with ideas, take risks by themselves, and learn from their own successes and failures. However, school work is typically defined by well-structured problems of relatively short duration and with fixed goals. Under

[2] Bereiter, C. and Scardamalia, M. (1993). *Surpassing ourselves: an inquiry into the nature and implications of expertise*. Open Court

[3] Dunbar, K. N. (1995). How scientists really reason: Scientific reasoning in real-world laboratories. In Sternberg, R. J. and Davidson, J., editors, *The Nature of Insight*, number 1995 in Mechanisms of insight, chapter 11, pages 365–395. MIT Press

[4] Gardner, H. (1994). More on private intuitions and public symbol systems. *Creativity Research Journal*, 7(3-4):265–275

[5] de Groot, A. (1978). *Thought and Choice in Chess*. Mouton De Gruyter, The Hague, The Netherlands, 2nd edition

[6] Buchanan, R. (1992). Wicked Problems in Design Thinking. *Design Issues*, 8(2):5–21

[8] Bereiter, C. (2002). *Education and Mind in the Knowledge Age*. Lawrence Erlbaum

these circumstances, students have little experience dealing with risks, uncertainty, choice-making and long-stretches of work, which abound in solving real-world ill-defined problems.

Knowledge Building pedagogy aims to fundamentally transform education into a coherent enterprise of knowledge creation in the whole society[9]. Students in knowledge-building classrooms treat their ideas as improvable and work collaboratively to improve them to achieve specific knowledge goals. In this situation, evaluation of promisingness naturally becomes an important component in their knowledge-building discourse, as nascent ideas are usually in abundance and to make progress students need to make choices on which ideas to spend time on. To engage students in promisingness judgments on their own ideas, we adopted a design research methodology[10] and developed a "Promising Ideas tool" which was seamlessly integrated with Knowledge Forum[11]. Although promisingness judgments do not come easily for young students[12], pedagogical and technological interventions show promise in facilitating this process and eventually boosting knowledge advancement in students' knowledge-building discourse[13].

Beside design research, another strand of research is needed to unpack the developmental aspect of promisingness judgments to inform further refinement of interventions. This study represents our first effort in this area, which tackles the following questions. Firstly, can developmental patterns in promisingness judgments be identified based on selections of ideas according to specific criteria, such as "promising for making a scientific breakthrough"? Second, if there is a developmental trajectory of such promisingness judgments, how is it be related to domain knowledge in relevant subject matter?

Methods

Instruments

A survey was designed to answer the research questions. The main goal of this survey was to engage participants in evaluating promisingness of a list of ideas in *optics*—a physics topic that has been widely discussed in science education literature. The survey contained three sections: (a) a demographics section, (b) an idea evaluation section, and (c) a quiz section. The demographics section collected relevant demographic information including age, institutions, and level of expertise in physics. The quiz section was composed of 10 multiple choice items adapted from an elementary science test, which was concerned with general content knowledge of optics, such as "how light travels," "shadow," "reflection" and "refraction".

[9] Scardamalia, M. and Bereiter, C. (2003). Knowledge building. In Guthrie, J. W., editor, *Encyclopedia of education*, volume 17, pages 1370–1373. Macmillan Reference, New York, NY, 2 edition

[10] Collins, A., Joseph, D., and Bielaczyc, K. (2004). Design Research: Theoretical and Methodological Issues. *Journal of the Learning Sciences*, 13(1):15–42

[11] Chen, B., Resendes, M., Chuy, M., Tarchi, C., Carl, B., and Scardamalia, M. (2011b). Identificare, selezionare e sviluppare le idee promettenti nel Knowledge Building. *QWERTY - Interdisciplinary Journal of Technology, Culture and Education*, 6(2):224–241

[12] Chen, B., Chuy, M., Resendes, M., Scardamalia, M., and Bereiter, C. (2011a). Evaluation by Grade 5 and 6 Students of the Promisingness of Ideas in Knowledge-Building Discourse. In Spada, H., Stahl, G., Miyake, N., and Law, N., editors, *Connecting Computer-Supported Collaborative Learning to Policy and Practice: CSCL2011 Conference Proceedings. Volume II - Short Papers and Posters*, number 2008, pages 571–575. International Society of the Learning Sciences

[13] Chen, B., Scardamalia, M., Resendes, M., Chuy, M., and Bereiter, C. (2012). Students' intuitive understanding of promisingness and promisingness judgments to facilitate knowledge advancement. In van Aalst, J., Thompson, K., Jacobson, M. J., and Reimann, P., editors, *The future of learning: Proceedings of the 10th international conference of the learning sciences (ICLS 2012) – Volume 1, Full Papers*, pages 111–118, Sydney, Australia. ISLS

The "idea evaluation" section contained 16 carefully drafted ideas about reflection and refraction in optics selected from various types of materials, including curriculum standards, misconception literature, and newest scientific reports. Some of them were rewritten to be comprehensible for young students and to ensure judgments of promisingness cannot be made from the surface.

Evaluation of ideas was based on a tentative promisingness scale composed of four levels including: "1-Not quite right", "2-Simple fact", "3-Deep understanding", and "4-Possible scientific breakthrough". Level 1 and 2 represent two levels not on the promisingness "spectrum." Ideas on Level 1 were classic misconceptions in optics. Comparing to items in the quiz, they focused on reflection and refraction and could provide a more specific measure of domain knowledge. Ideas on Level 2 are simple facts mainly from textbooks or curriculum standards, presented without any further explanation. Ideas on Level 3 are ideas picked from textbooks or other authoritative sources containing a scientific explanation of an optical phenomenon. Level 4 were designed to embody the most straightforward criterion of promisingness, i.e. "promising for making a scientific breakthrough." Idea at this level come from scholarly articles about new findings in this field. Example of items are presented in Figure 1.

Levels	**Examples**
1-Not quite right	The convergent lens increases the speed and energy of the light. That's why it can make a piece of paper on fire.
2-Simple fact	When light hits a mirror, the angle of incidence equals the angle of reflection.
3-Deep understanding	Refraction is the bending of light when it passes from one kind of material into another. Because light travels at a different speed in different materials, it must change speeds at the boundary between two materials.
4-Possible breakthrough	Three-dimensional solar cells on solar panels can trap sunlight inside a structure, making light bounce around until it is converted into electricity. This makes it more efficient in converting sunlight to electricity.

Table 1: Means and standard deviations of age in each group

Participants

Students from Grade 4 (n = 20), Grade 5 (n = 21), and Grade 7 (n = 23) were recruited to participate in this study. The Grade 4 class just learned the optics unit in the past school year, while the Grade 5 students did not learn that unit when they were in Grade 4. Grade 7 students studied optics two years ago when they were in Grade 5 and a few new students in the class had not learned optics before. Adults (n = 5) and domain experts (n = 4) were also invited to participate. The adult group was made up of adults who did not have a post-secondary physics background, while domain experts were faculty members or doctoral students from a physics department.

Data Analysis

All complete responses from participants were collected for analysis. First of all, each participant's performance of evaluating ideas on each level was measured by counting number of matches with the predefined level (or the "correct" answer). Given numbers of items were not equal for each level[14], the final score of each level ($score_1$ - $score_4$[15]) was weighted by the number of items. Since $score_1$ was treated as a measure of domain knowledge, a sum of scores of other three levels, $score_{sum}$, was computed to measure one's overall performance in promisingness judgments. However, measuring by direct matches would left out ordinal information of levels in a participant's judgments. To provide a more accurate measure that can capture ordinal information, we calculated Krippendorff's (1980) alpha coefficient of agreement, $score_\alpha$, between one's responses with the "correct" answers. As for the quiz unit, each participant was assigned a score (from 0 to 10) based on how many questions she had correctly answered. Moreover, we also collected each participant's group membership, age and completion time for analysis.

[14] Level 1 to 4 had 5, 3, 3 and 5 ideas respectively.

[15] $score_i = \frac{hits}{n} \in [0,1]$, where i denotes the level, $hits$ denotes the number of matches at this level, and n denotes the number of ideas at this level.

To inspect possible developmental patterns, we compared mean scores of judgment performance on each idea level, overall performance of the whole evaluation task, and quiz scores across different age groups. One-way analysis of variance (ANOVA) was performed to achieve this goal. When interpreting results of ANOVA, detailed analysis of judgments on each idea was conducted when necessary. To investigate relationships between participant performance on idea evaluation and other variables such as age, domain knowledge and completion time, bivariate correlations were analyzed.

Videos of students' reflective discussion after completing the survey were collected from Grade 4 and 5 classes. For participants from other groups, we included an open-ended question in this survey to collect their reflections or suggestions on this survey. Video discussions and responses to the open-ended question were qualitatively analyzed to inform our interpretation of statistical analysis.

Results

Completion Time, Age and Content Knowledge

Before investigating possible developmental patterns of promisingness judgments, we first compared the means of time spent on the survey among all groups, to ensure all responses were valid and comparisons were meaningful. ANOVA did not find any significant difference across groups, $F(4, 71) = 2.29$, $p = .07$. However, two outliers in Grade 7, who spent less than 250 seconds on the survey, were identified. Because completing this survey was a cognitively demanding task and participants normally needed a considerable amount of time to complete ($M = 779.61$, $SD = 332.4$), these two outliers were left out in further analysis.

Second, to get a sense how participants varied in terms of age, we compared means of age across different groups. Means and standard deviations are presented in Table 2. Because the age groups between Grade 7 and adults were not represented in this study, interpretation of any developmental pattern in this study needs to be cautious.

Table 2: Means and standard deviations of age in each group

Groups	*Means*	*SD*
Grade 4	9.45	0.51
Grade 5	10.43	0.51
Grade 7	13.76	0.54
Adult non-experts	28.8	5.26
Experts	37.75	10.63

To compare content knowledge among groups, one-way ANOVA tests were performed on two variables, including quiz score (*quiz*) and evaluation performance on "Level 1-Not quite right" items ($score_1$). Results showed significant difference in both *quiz* ($F(4, 66) = 4.58$, $p < .001$) and $score_1$ ($F(4, 66) = 16.02$, $p < .001$). Further post-hoc tests indicated two adult groups (i.e., adult non-experts and experts) showed significantly more advanced domain knowledge than three student groups, while all other group contrasts were nonsignificant, except for the comparison of $score_1$ between adult non-experts and experts. More detailed analysis on each item revealed young students from Grade 4, 5 and 7 were making a lot of mistakes in evaluating ideas on Level 1, misjudging them as "simple fact" or "deep understanding", while experts could recognize all of them as misconceptions. These results implied adults generally had better content knowledge of optics than young students, although adult non-experts held some classic misconceptions in this domain.

Developmental Patterns of Promisingness Judgments

The central research questions in this study were how participants performed differently in the idea evaluation task and how was their performance related to age and domain knowledge. To answer the first question, we analyzed group difference of performance in evaluating Level 2-4 ideas, i.e., $score_{2-4}$. For each level, a score was calculated for each participants based on the number of hits compared to the "correct" answer weighted by the number of items at that level.

Each participant's overall performance was measured by the sum of scores on each level ($score_{sum}$) and Krippendorff's (1980) alpha coefficient of agreement ($score_{\alpha}$). One-way ANOVA tests were performed on these measures. Results indicated mean scores of $score_2$ ($F(4, 66) = 1.01$, $p = .41$) and $score_3$ ($F(4, 66) = 0.93$, $p = .45$) were not significantly different across groups (see Figure 1). For ideas on Level 2, further analysis found a number of young students had difficulty distinguishing ideas on this level (i.e., "simple facts") from "deep understanding" or "possible scientific breakthrough." One possible cause of this result was their limitation in domain knowledge; because they had limited content knowledge, they tended to recognize facts beyond their knowledge as deep understanding or scientific breakthroughs. For ideas on Level 3, responses were blended, as participants in all groups tended to confound them with "simple fact." Further analysis found experts were likely to classify ideas on Level 3 (as well as Level 2) to lower levels for subtle, rhetorical reasons, while participants in other groups were inclined to classify "deep understanding" they knew well as "simple fact." Overall, no group difference was found on idea evaluation performance on the levels of "2-Simple fact" and "3-Deep understanding," and judgments on these two levels could either benefit from or be disburbed by personal domain knowledge.

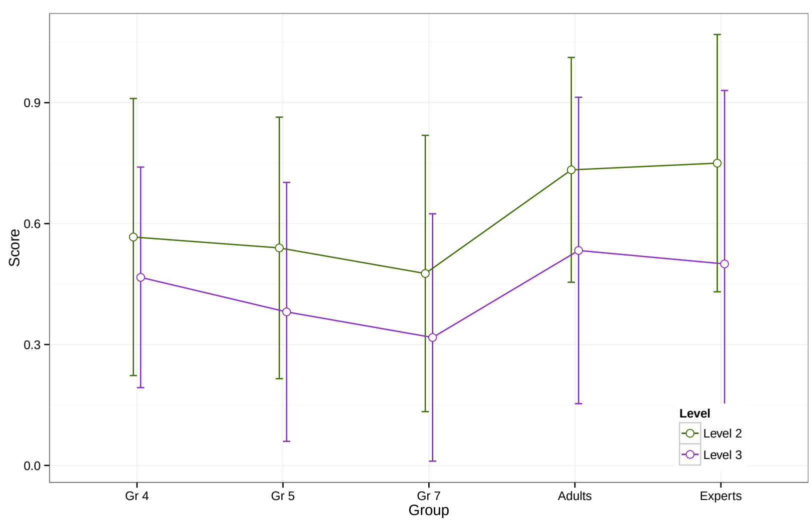

Figure 1: Mean scores of level 2 and 3 for each group.

Evaluation on ideas at Level 4—"possible scientific breakthrough"—was of the most interest in this study as this level represents one most straightforward criterion of promisingness. One-way ANOVA identified significantly different performance of evaluating ideas at this level among five groups, $F(4, 66) = 3.73$, $p < .01$. Means and standard deviations are presented in Figure 2. Further post-hoc tests found adult non-experts did significantly better than Grade 5 who had not study optics before (Cohen's $d = 0.33$, $p < .05$), while all other comparisons were nonsignificant on the significance level of .05. In other words, performance of Grade 4 and 7 students, who had studied optics before the study, was not significantly inferior than adult groups on this task. Further analysis of their reflective discussion showed some students had a quite reliable sense of promisingness and it could be mediated by content knowledge in action. For instance, when discussing a breakthrough idea about the mechanism of detecting cancer with laser, one Grade 4 student explained the reason he chose Level 4 was "*because they are closer to discover how to distroy cancer, and that is a scientific breakthrough*," whereas another student thought "*it could also be a thorough explanation because ... I sort of have already knew because my dad is a doctor and I go to hospital a lot*." Interestingly, another student who chose "not quite right" for that idea explained that she "*did not get it and thought it meant the microscope kills the cancer cells*." As for experts, a closer look at evaluation on each idea as well as their qualitative feedback found experts brought in alternative theories that challenges those "scientific breakthroughs" and deemed them incorrect. For example, one expert admitted his rich domain knowledge was confusing his judgment on one breakthough finding about invisibility cloak: "*With respect to the harry potter thing (i.e., invisibility cloak), suffice to say that it's a little different from just interference. But I think my further knowledge in some parts is confusing the issue.*" For this reason, rich domain knowledge of experts actually was making things more complex and could prevent them from making "correct" promisingness judgments. That means, experts were not unable to identify promising breakthroughs but tended to see those new breakthroughs in different ways depending on their favored domain perspectives. Thus, it becomes not surprising that their mean score on promisingness judgments was lower than adult non-experts. Overall, these results indicated young students, as long as they have some basic knowledge, could have a not significantly inferior sense of promisingness at least than adult non-experts.

Correlation analysis was further performed to investigate the relationship between different variables (see Table 3). Results indicated that general performance on the idea evaluation task measured by $score_\alpha$ was significantly correlated with age ($r = .32$, $p < .01$) and domain knowledge measured by quiz ($r = .23$, $p < .05$). More specifically, the capability of promisingness judgments in the context of scientific breakthroughs was also significantly related to both age ($r = .38$, $p < .01$) and domain knowledge measured by the ability to identify classic misconceptions ($r = .24$, $p < .05$). However, since

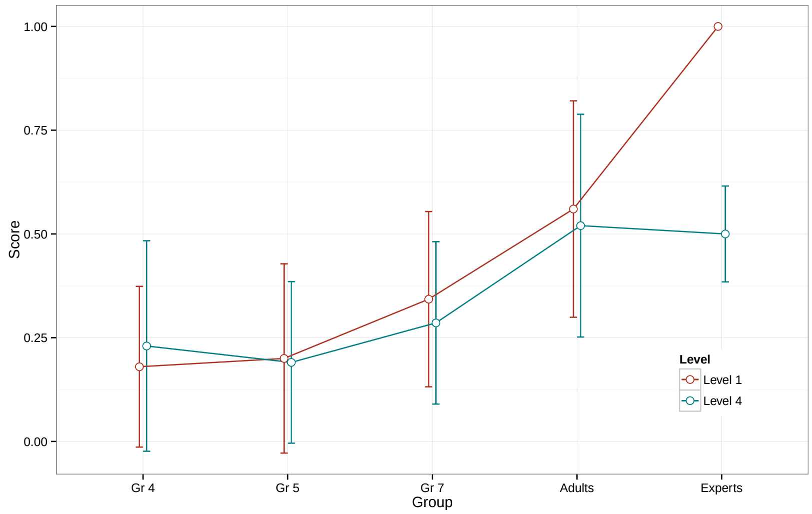

Figure 2: Mean scores of level 1 and 4 for each group.

correlation analysis showed an even stronger correlation between age and two domain knowledge variables (r = .37, p < .01 and r = .64, p < .001) and the age groups between Grade 7 and adults were missing, it was not clear which factor—developmental difference or domain knowledge—had a greater impact on promisingness judgments. Future research with revised instruments and participation from broader age groups will try to tackle this issue.

Variables	1	2	3	4	5	6
1. age	-					
2. quiz	.37**	-				
3. $score_1$	.64***	.27*	-			
4. $score_3$	.15	.09	-.06	-		
5. $score_4$	.38**	.14	.24*	.01	-	
6. $score_\alpha$	.32**	.23*	.11	.27*	.73***	-

Table 3: Correlations between Measures. *Note:* * p < .05. ** p < .01. *** p < .001.

Conclusions

Promisingness judgments are thought to be the heart of effective, creative action, and improved by immersion in progressive problem solving[16]. While inventors usually need to make choice involves risk and uncertainty when dealing with ill-defined problems, school work is typically defined by well-defined problems of relatively short duration with fixed goals and continuous, clear feedback. This study is part of our ongoing research program that aims to increase students' innovative capacity by accustoming them to thinking in terms

[16] Bereiter, C. and Scardamalia, M. (1993). *Surpassing ourselves: an inquiry into the nature and implications of expertise*. Open Court

of promising ideas and maximizing possibilities of conceptual breakthrough when working towards knowledge-building goals. In this preliminary study, we found that domain knowledge would generally shed light on promisingness judgments by people from different age groups. However, as long as young students adopt some basic understanding of a subject matter, they could have a considerable sense of promisingness in that domain. More in-depth analysis found performance of promisingness judgments correlated with both age and domain knowledge. Future research is clearly needed to achieve a clearer understanding of the developmental aspect of promisingness judgments, which will inform our design-based research to facilitate students' promisingness judgments in the future.

References

Bereiter, C. (2002). *Education and Mind in the Knowledge Age*. Lawrence Erlbaum.

Bereiter, C. and Scardamalia, M. (1993). *Surpassing ourselves: an inquiry into the nature and implications of expertise*. Open Court.

Buchanan, R. (1992). Wicked Problems in Design Thinking. *Design Issues*, 8(2):5–21.

Chen, B., Chuy, M., Resendes, M., Scardamalia, M., and Bereiter, C. (2011a). Evaluation by Grade 5 and 6 Students of the Promisingness of Ideas in Knowledge-Building Discourse. In Spada, H., Stahl, G., Miyake, N., and Law, N., editors, *Connecting Computer-Supported Collaborative Learning to Policy and Practice: CSCL2011 Conference Proceedings. Volume II - Short Papers and Posters*, number 2008, pages 571–575. International Society of the Learning Sciences.

Chen, B., Resendes, M., Chuy, M., Tarchi, C., Carl, B., and Scardamalia, M. (2011b). Identificare, selezionare e sviluppare le idee promettenti nel Knowledge Building. *QWERTY - Interdisciplinary Journal of Technology, Culture and Education*, 6(2):224–241.

Chen, B., Scardamalia, M., Resendes, M., Chuy, M., and Bereiter, C. (2012). Students' intuitive understanding of promisingness and promisingness judgments to facilitate knowledge advancement. In van Aalst, J., Thompson, K., Jacobson, M. J., and Reimann, P., editors, *The future of learning: Proceedings of the 10th international conference of the learning sciences (ICLS 2012) – Volume 1, Full Papers*, pages 111–118, Sydney, Australia. ISLS.

Collins, A., Joseph, D., and Bielaczyc, K. (2004). Design Research:

Theoretical and Methodological Issues. *Journal of the Learning Sciences*, 13(1):15–42.

de Groot, A. (1978). *Thought and Choice in Chess*. Mouton De Gruyter, The Hague, The Netherlands, 2nd edition.

Dunbar, K. N. (1995). How scientists really reason: Scientific reasoning in real-world laboratories. In Sternberg, R. J. and Davidson, J., editors, *The Nature of Insight*, number 1995 in Mechanisms of insight, chapter 11, pages 365–395. MIT Press.

Gardner, H. (1994). More on private intuitions and public symbol systems. *Creativity Research Journal*, 7(3-4):265–275.

Johnson, J. (2006). Cognitive modeling of decision making in sports. *Psychology of Sport and Exercise*, 7(6):631–652.

Scardamalia, M. and Bereiter, C. (2003). Knowledge building. In Guthrie, J. W., editor, *Encyclopedia of education*, volume 17, pages 1370–1373. Macmillan Reference, New York, NY, 2 edition.