Beyond Navel Gazing: The Evidence Base for Employing Reflective and Metacognitive Practices in our Teaching

Sarah Leupen
University of Maryland

Follow this and additional works at: https://digitalcommons.georgiasouthern.edu/sotlcommons

Part of the Curriculum and Instruction Commons, Educational Assessment, Evaluation, and Research Commons, Educational Methods Commons, Higher Education Commons, and the Social and Philosophical Foundations of Education Commons

Recommended Citation
https://digitalcommons.georgiasouthern.edu/sotlcommons/SoTL/2016/7

This presentation (open access) is brought to you for free and open access by the Conferences & Events at Digital Commons@Georgia Southern. It has been accepted for inclusion in SoTL Commons Conference by an authorized administrator of Digital Commons@Georgia Southern. For more information, please contact digitalcommons@georgiasouthern.edu.
Beyond Navel Gazing

The Evidence Base for Employing Reflective and Metacognitive Practices in our Teaching

Sarah Leupen
UMBC
Traxoline is a new form of zionter. It is montillated in Ceristanna. The Ceristannians gristerlate large amounts of fevon and then bracter it to quasel traxoline. Traxoline may be one of our most lukized snezlaus in the future.
Traxoline is a new form of zionter. It is montillated in Ceristanna. The Ceristannians gristerlate large amounts of fevon and then bracter it to quasel traxoline. Traxoline may be one of our most lukized snezlaus in the future.

How is trazoline quaselled?

a) Via gristerlation
b) Via bracteration
c) Via lukization
d) Via montillation
Metacognition: The act of thinking about one’s own thought processes

Reflection: Conscious exploration of one’s own experiences
Everyone thinks they got a B!

Our key interest was how perceptions of that ability converged with actual ability. Specifically, we wanted to discover whether those who did poorly on our measure would recognize the low quality of their performance. Would they recognize it or would they be unaware?

4.1.1. Method

Participants. Participants were 65 Cornell University undergraduates from a variety of courses in psychology who earned extra credit for their participation.

Materials. We created a 30-item questionnaire made up of jokes we felt were of varying comedic value. Jokes were taken from Woody Allen (1975), Al Franklin (1992), and a book of "really silly" pet jokes by Jeff Rovin (1996). To assess joke quality, we contacted several professional comedians via electronic mail and asked them to rate each joke on a scale ranging from 1 (not at all funny) to 11 (very funny). Eight comedians responded to our request (Bob Crawford, Costaki Economopoulos, Paul Frisbie, Kathleen Madigan, Ann Rose, Allan Sitterson, David Spark, and Dan St. Paul). Although the ratings provided by the eight comedians were moderately reliable ($a = .72$), an analysis of interrater correlations found that one (and only one) comedian's ratings failed to correlate positively with the others (mean $r = -.09$). We thus excluded this comedian's ratings in our calculation of the humor value of each joke, yielding a final $a$ of .76.

Expert ratings revealed that jokes ranged from the not so funny (e.g., "Question: What is big as a man, but weighs nothing? Answer: His shadow." Mean expert rating = 1.3) to the very funny (e.g., "If a kid asks where rain comes from, I think a cute thing to tell him is 'God is crying.' And if he asks why God is crying, another cute thing to tell him is 'probably because of something you did.'" Mean expert rating = 9.6).

Procedure. Participants rated each joke on the same 11-point scale used by the comedians. Afterward, participants compared their "ability to recognize what's funny" with that of the average Cornell student by providing a percentile ranking. In this and in all subsequent studies, we explained that percentile rankings could range from 0 (I'm at the very bottom) to 50 (I'm exactly average) to 99 (I'm at the very top).

4.1.2. Results and Discussion

Gender failed to qualify any results in this or any of the studies reported in this article, and thus receives no further mention.

Our first prediction was that participants overall would overestimate their ability to tell what is funny relative to their peers. To find out whether this was the case, we first assigned each participant a percentile rank based on the extent to which his or her joke ratings correlated with the ratings provided by our panel of professionals (with higher correlations corresponding to better performance). On average, participants put their ability to recognize what is funny in the 66th percentile, which exceeded the actual mean percentile (50, by definition) by 16 percentile points, one-sample $t(64) = 7.02$, $p < .0001$. This overestimation occurred even though self-ratings of ability were significantly correlated with our measure of actual ability, $r(63) = .39$, $p < .001$.

Our main focus, however, is on the perceptions of relatively "incompetent" participants, which we defined as those whose test score fell in the bottom quartile ($n = 16$). As Figure 1 depicts, these participants grossly overestimated their ability relative to their peers. Whereas their actual performance fell in the 12th percentile, they put themselves in the 58th percentile. These estimates were not only higher than the ranking they actually achieved, paired $t(15) = 10.33$, $p < .0001$, but were also marginally higher than a ranking of "average" (i.e., the 50th percentile), one-sample $t(15) = 1.96$, $p < .07$. That is, even participants in the bottom quarter of the distribution tended to feel that they were better than average.

As Figure 1 illustrates, participants in other quartiles did not overestimate their ability to the same degree. Indeed, those in the top quartile actually underestimated their ability relative to their peers, paired $t(15) = -2.20$, $p < .05$.

Figure 1. Perceived ability to recognize humor as a function of actual test performance (Study 1). Copyright © 2009 SciRes. Kruger and Dunning, Psychology, 2009, 1, 30-46
Unskilled and Unaware of It: How Difficulties in Recognizing One's Own Incompetence Lead to Inflated Self-Assessments

Justin KRUGER, David DUNNING

People tend to hold overly favorable views of their abilities in many social and intellectual domains. The authors suggest that this overestimation occurs, in part, because people who are unskilled in these domains suffer a dual burden: Not only do these people reach erroneous conclusions and make unfortunate choices, but their incompetence robs them of the metacognitive ability to realize it. Across 4 studies, the authors found that participants scoring in the bottom quartile on tests of humor, grammar, and logic grossly overestimated their test performance and ability. Although their test scores put them in the 12th percentile, they estimated themselves to be in the 62nd. Several analyses linked this misattribution to deficits in metacognitive skill, or the capacity to distinguish accuracy from error. Paradoxically, improving the skills of participants, and thus increasing their metacognitive competence, helped them recognize the limitations of their abilities.

Keywords: Inflated Self-Assessments, metacognitive competence

In 1995, McArthur Wheeler walked into two Pittsburgh banks and robbed them in broad daylight, with no visible attempt at disguise. He was arrested later that night, less than an hour after videotapes of him taken from surveillance cameras were broadcast on the 11 o'clock news. When police later showed him the surveillance tapes, Mr. Wheeler stared in incredulity. "But I wore the juice," he mumbled. Apparently, Mr. Wheeler was under the impression that rubbing one's face with lemon juice rendered it invisible to videotape cameras (Fuocco, 1996).

We bring up the unfortunate affairs of Mr. Wheeler to make three points. The first two are noncontroversial. First, in many domains in life, success and satisfaction depend on knowledge, wisdom, or savvy in knowing which rules to follow and which strategies to pursue. This is true not only for committing crimes, but also for many tasks in the social and intellectual domains, such as promoting effective leadership, raising children, constructing a solid logical argument, or designing a rigorous psychological study. Second, people differ widely in the knowledge and strategies they apply in these domains (Dunning, Meyerowitz, & Holzberg, 1989; Dunning, Perie, & Story, 1991; Story & Dunning, 1998), with varying levels of success. Some of the knowledge and theories that people apply to their actions are sound and meet with favorable results. Others, like the lemon juice hypothesis of McArthur Wheeler, are imperfect at best and wrong-headed, incompetent, or dysfunctional at worst.

Perhaps more controversial is the third point, the one that is the focus of this article. We argue that when people are incompetent in the strategies they adopt to achieve success and satisfaction, they suffer a dual burden: Not only do they reach erroneous conclusions and make unfortunate choices, but their incompetence robs them of the ability to realize it. Instead, like Mr. Wheeler, they are left with the mistaken impression that they are doing just fine. As Miller (1993) perceptively observed in the quote that opens this article, and as Charles Darwin (1871) sagely noted over a century ago, "ignorance more frequently begets confidence than does knowledge" (p. 3). In essence, we argue that the skills that engender competence in a particular domain are often the very same skills necessary to evaluate competence in that domain—one's own or anyone else's. Because of this, incompetent individuals lack what cognitive psychologists variously term metacognition (Everson & Tobias, 1998),
Which of the following is the most important ingredient for successful learning?

a) The intention and desire to learn
b) Paying close attention to the material as you study
c) The time you spend studying
d) Matching your own learning style as you study
e) What you think about while studying
Expecting to teach enhances learning and organization of knowledge in free recall of text passages

John F. Nestojko · Dung C. Bui · Nate Kornell · Elizabeth Ligon Bjork

UCLA students all read the same passage about the Crimean War

Some expected to take a test on it, and others expected to be asked to teach it to others

Actually, they all just took a test
Expecting to teach enhances learning and organization of knowledge in free recall of text passages

John F. Nestojko · Dung C. Bui · Nate Kornell · Elizabeth Ligon Bjork

![Graph showing proportion of SA questions answered under different expectancy conditions]
1. Teach students what we actually know about learning these days!

![Bloom's Taxonomy Diagram]

- **Remembering**: Retrieving, recognizing and recalling knowledge from long-term memory.
- **Understanding**: Interpreting, exemplifying, classifying, summarizing, inferring, comparing and explaining.
- **Applying**: Carrying out a procedure through executing or implementing.
- **Analyzing**: Breaking material into parts and relating them to one another and to an overall structure.
- **Evaluating**: Making judgments based on criteria and standards through checking and critiquing.
- **Creating**: Putting elements together, reorganizing elements into a new pattern or structure through generating, planning or producing.

---

**Figure 1.** S. Taxonomy

Metacognition: An Effective Tool to Promote Success in College Science Learning

By Ningfeng Zhao, Jeffrey G. Wardeska, Saundra Y. McGuire, and Elzbieta Cook

[Diagram of the revised Bloom's Taxonomy]

- **Remembering**: Retrieving, recognizing and recalling knowledge from long-term memory.
- **Understanding**: Interpreting, exemplifying, classifying, summarizing, inferring, comparing and explaining.
- **Applying**: Carrying out a procedure through executing or implementing.
- **Analyzing**: Breaking material into parts and relating them to one another and to an overall structure.
- **Evaluating**: Making judgments based on criteria and standards through checking and critiquing.
- **Creating**: Putting elements together, reorganizing elements into a new pattern or structure through generating, planning or producing.

The revised Bloom's Taxonomy represents a hierarchy of learning levels.
Metacognition: An Effective Tool to Promote Success in College Science Learning

By Ningfeng Zhao, Jeffrey G. Wardeska, Saundra Y. McGuire, and Elzbieta Cook

- Used a whole class period to train students in effective learning techniques
- Explained Bloom's taxonomy using test questions from the first exam to illustrate
- Described metacognitive study strategies
Metacognition: An Effective Tool to Promote Success in College Science Learning

By Ningfeng Zhao, Jeffrey G. Wardeska, Saundra Y. McGuire, and Elzbieta Cook

FIGURE 1
Student exam performance for the fall semesters.
The Chemistry Self-Concept Inventory is an effective instrument with proven reliability (Bauer, 2005). In fall 2011, the postsurvey showed higher means in almost all categories, with the largest improvement in chemistry self-concept (Table 3 in blue), suggesting that through metacognition and implementation of effective learning strategies, students improved in chemistry learners. The postsurvey in spring 2012 showed significant improvements in math self-concept and academic enjoyment from paired sample $t$-tests (Table 3 in blue). The intensive involvement of mathematics in General Chemistry II, such as kinetics and equilibrium calculations, might have contributed here. Again, through metacognition and effective learning, students reported improved academic enjoyment. Somewhat surprisingly, the postsurvey in spring 2012 showed a decrease in the chemistry self-concept (Table 3 in red). Although this might be related to the breadth and depth of topics involved in the full-year general chemistry courses. Considering the much more intensive study each of these topics requires, students might have realized the limits of their chemistry knowledge and/or become more critical of themselves.

Data collected from the 11 dual participants (those who were also participants in fall 2011) in spring 2012 were analyzed separately. Not surprisingly, they had a better understanding of the required levels of learning in college, as well as generally higher means in both presurveys when compared with their peers in spring 2012, who had not learned the strategies in fall 2011. For the postsurvey of the Effective Learning Strategies Survey, they also had generally lower means for Statements 3–6 and higher means for Statements 7–12 (Table 2). For the Chemistry Self-Concept Inventory, there were similar improvements in math, academic, and creativity self-concepts, and decline in chemistry self-concept (Table 3).

The first author taught General Chemistry I in fall 2011, 2010, and 2009, and the second author taught General Chemistry II in spring 2012, 2011, and 2010 with the same curriculum and schedule. In this pilot study, comparisons were not established for equivalence of students or exams from year to year. Nevertheless, exam questions were selected from the same test bank (http://www.wileyplus.com) with already predetermined levels.
Do Metacognitive Exercises Improve Critical Thinking?

Figure 3.1  Student learning goals

By the end of this exercise, you should be able to:

- Describe a strategy for tackling scientific questions.
- Describe the structure of answers to exam questions that assess critical thinking, considering the relative quality of different sample answers. This goal includes detailed analysis of the components of a complete answer, including content knowledge and critical thinking.
- Apply your improved understanding by answering another scientific problem on your own and with less guidance.
- Analyze your own answer, as well as those of your peers, in order to determine the strengths and weaknesses in your critical-thinking skills.

Figure 3.2  An example question

When a genetic disorder is diagnosed in a family, family members often want to know the likelihood that they or their children will develop the condition. One important factor that influences a person's chances of developing a genetic condition is how the condition is inherited (i.e., whether it is autosomal or sex-linked, dominant or recessive, simple or codominant).

![Family tree diagram]

A. In examining the pedigree, what mode of inheritance best describes the trait indicated by the dark symbols? For full credit, you must name a mode of inheritance, explain what you think the most likely phenotype (affected or unaffected) and genotype of individual #1 would be, and explain why. (Hint: to make an argument that one mode of inheritance is more likely than others, you will need to discuss the other modes of inheritance and explain why they are less likely.)
Metacognitive Exercises Improve Critical Thinking!

Figure 3.11 Discussion starters to help students reflect on the classroom exercise

- What do you think of the questions and answers you’ve seen today? Are they what you expected? How different? How similar?
- In the answers you looked at today, were there different ways to arrive at a moderate-quality answer? A high-quality answer?
- How did you approach answering the questions you were given in class today? Could you have found the answers in your textbook? What are some of the thinking skills you had to use while you were answering the questions? Did you need to recall info, explain, apply, analyze, or evaluate?
- What are some study strategies you could use to prepare for these types of questions?

Table 3.1 A one-tailed, $t$-test comparison of 35 paired pre- and posttests

<table>
<thead>
<tr>
<th></th>
<th>Mean pretest score (%)</th>
<th>Mean posttest score (%)</th>
<th>$t$ statistic</th>
<th>$p$ value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Content</td>
<td>86</td>
<td>81</td>
<td>0.90</td>
<td>0.19</td>
</tr>
<tr>
<td>Critical thinking</td>
<td>75</td>
<td>86</td>
<td>-2.15</td>
<td>0.01</td>
</tr>
</tbody>
</table>
2. Use Learning Wrappers
2. Use Learning Wrappers

33-111 Physics Post-Exam Reflection

As with the first exam, this activity is designed to give you a chance to reflect on your exam performance and, more importantly, on the effectiveness of your exam preparation. Again, please answer the questions sincerely. Your responses will be collected to inform the instructional team; they will have no impact on your grade.

1. Approximately how much time did you spend preparing for this exam? _______

2. What percentage of your test-preparation time was spent in each of these activities?
   a. Reading textbook section(s) for the first time _______
   b. Re-reading textbook section(s) _______
   c. Reviewing homework solutions _______
   d. Solving problems for practice _______
   e. Reviewing your own notes _______
   f. Reviewing materials from blackboard _______
      (What materials? _____________________)
   g. Other _______
      (Please specify: ______________________)

3. What aspect(s) of your preparation for this exam seemed different from your exam 1 preparation? Did these changes have any effect?

4. Now that you have looked over your graded exam, estimate the percentage of points you lost due to each of the following (make sure the percentages add up to 100):
   a. Trouble with vectors and vector notation _______
   b. Algebra or arithmetic errors _______
   c. Problem with force-body diagram _______
   d. Lack of understanding of the concept _______
   e. Not knowing how to approach the problem _______
   f. Careless mistakes _______
   g. Other _______
      (Please specify: ______________________)

PLEASE CONTINUE ON THE BACK IF YOU NEED MORE ROOM.
Exam Wrappers Help Students Identify Effective Strategies

Figure 2.3 Students’ change in rating of “effective” strategies as a function of exam wrapper dose

Lovett, Make Exams Worth More than the Grade, in: Using Reflection and Metacognition to Improve Student Learning, Stylus 2013
... And Ineffective Strategies!

**Figure 2.4** Students’ change in rating of “less effective” strategies as a function of exam wrapper dose

Lovett, Make Exams Worth More than the Grade, in: Using Reflection and Metacognition to Improve Student Learning, Stylus 2013
5. Based on your answers above, if you are not fully satisfied with your test performance, list 2-3 things you will do differently in preparing for Test 2.

- go over more concept based things instead of memorization
- effect of phospho channels etc.
- study more with a group than alone.

5. Based on your answers above, if you are not fully satisfied with your test performance, list 2-3 things you will do differently in preparing for Test 2.

- Practice hypothetical situations such as injecting TRH & predicting effects

5. Based on your answers above, if you are not fully satisfied with your test performance, list 2-3 things you will do differently in preparing for Test 2.

1) Begin studying sooner, trying to stay ahead of class
2) Try to teach someone the learning goals
Wrappers: Not just for exams any more!

Our understanding of intelligence, learning, the brain has exploded in the last 30 years. We’ve discovered multiple types of intelligence, neural plasticity and realized that you can teach an old dog new tricks. One of the most important discoveries is that Dewey was right: metacognition or “thinking about your thinking” can improve learning and students with math anxiety who write about it before the exam, perform better than those who just take the exam (Ramirez & Beilock, 2011). Two different studies, one led by Marsha C. Lovett, (2013) and the other by Mary-Ann Winkelmes (2013), demonstrate how metacognition and reflection can be used to improve study habits, exam performance, and ultimately, new, easy and fast techniques to creating more critical thinkers.
The single *most* effective study strategy for students is:

a) taking a practice test  
b) rereading the textbook  
c) teaching someone else  
d) working example problems
3. Have Students Teach Each Other
3. Have Students Teach Each Other

Article

Is Peer Interaction Necessary for Optimal Active Learning?

Debra L. Linton,* Jan Keith Farmer,† and Ernie Peterson*

*Department of Biology, Central Michigan University, Mount Pleasant, MI 48858; †Department of Biology, Schoolcraft College, Livonia, MI 48152

• Same instructor
• Same semester
• Same instructional materials
• Same tests and other assessments

→ Active-learning exercises were completed either individually or in cooperative groups
Performance on lower-level Bloom questions

![Bar chart showing performance on lower-level Bloom questions](chart.png)
Performance on higher-level Bloom questions

![Bar chart showing performance on exams with error bars.](chart.png)
Why do students do better on a clicker question after discussing it with their peers?

a) Students with greater understanding tell the other students what the right answer is
b) Students with greater understanding help the other students understand the concept
c) Discussion helps students with poorer understanding notice the errors in their own thinking and correct them
Why Peer Discussion Improves Student Performance on In-Class Concept Questions


2 JANUARY 2009 VOL 323 SCIENCE www.sciencemag.org
Why Peer Discussion Improves Student Performance on In-Class Concept Questions


![Bar chart showing percentage of correct answers for Q1, Q1ad, and Q2 questions with individual vote and revote after peer discussion.]
Why Peer Discussion Improves Student Performance on In-Class Concept Questions

M. K. Smith,1* W. B. Wood,1 W. K. Adams,2 C. Wieman,2,3 J. K. Knight,1 N. Guild,1 T. T. Su1

Supporting Online Material

Fig. 1. The histograms of student answers were revealed after peer discussion; Q2: the isomorphic question Q2. Of the 32 questions in our 16 question pairs, 7 had 5 answer choices, 5 had 4, and 3 had 3 choices, and 1 had 2 choices.

Table 1. Data pooled from individual mean scores on Q1, Q1ad, Q2, and Q1ad* (%)

No significant improvement between these questions.
4. Use Team-Based Learning
The Team-Based Learning Sequence

4S Framework

- Significant Problem
- Same Problem
- Specific Choice
- Simultaneous Report

In the TBL classroom, the bulk of class time is spent having student teams solve, report, and discuss solutions to relevant, significant problems. Structuring the problems using TBL's 4S Framework lets you leverage the power of team processing without many of the problems that are inherent in other forms of small-group learning.

**Significant Problem**
Examples of Significant Problem
- A historian reconciles conflicting sources.
- A doctor decides the best course of action.
- A businessperson picks the best location for a business.
- A writer identifies the most powerful passage or best example.

You must use a significant, relevant problem that captures the interest of students. The quality of the problem ultimately controls the effectiveness of an application activity. Problems must require students to use course concepts to solve them.

**Same Problem**
Teams work on the same problem. This ensures the comparability of team solutions and this naturally acts as a potent discussion starter. Having students work on the same problem lets you create reporting opportunities for teams to defend, challenge, discuss, and examine each other's thinking and problem-solving process. Working on the same problem, ensures that students are interested in what other teams decided.

**Specific Choice**
Teams select the best choice from a limited list of options. This ensures that teams can easily compare their final decisions to the decisions of other teams. It is this comparability that drives the rich reporting discussion as teams examine and critique other teams' decisions and defend their own.

**Simultaneous Report**
Simultaneous reporting is most simply accomplished with holding up of a coloured card indicating a particular choice. When a team sees that another team has made a different decision, they naturally want to challenge the other teams' decision. In the ensuing conversation, the teams challenge each other and defend their own thinking. The reporting requires teams to articulate their thinking to other teams – putting their thoughts into words. This helps cognitively with the process of creating enduring, deep understanding. The feedback from their peers is immediate and focused on "how did you arrive at your decision" and not "which is the right answer."
- Linear Algebra taught using traditional lectures in Fall 2011 and Spring 2012
- Used Team-Based Learning in Fall 2012 and Spring 2013 (2 sections)
- Homework assignments, difficulty of quizzes and exams were held constant
People Prefer Electric Shocks to Being Alone With Their Thoughts

A new study finds we're not very good at entertaining ourselves.
5. Train Their Attention

Contemplative reading

Reflective writing
Don’s Reflective Writing Assignment

For this writing assignment students should print out and read Nicholas Carr’s chapter “The Juggler’s Brain” from The Shallows and Small and Vorgan’s “Your Brain is Evolving Right Now” locked away in a room with no distractions at all (no electronic media of any kind, no food, etc.). The student is allowed to take notes, but only with pen/pencil and paper. The paper should not be a review of the two readings, but rather the student should construct their own thesis (argument) and then use the two readings as support for your claims. The thesis can be reflective - based on your own thoughts about the process of reading and writing the assignment (you are encouraged to use words like “I”). Your argument should be inspired by the readings, not duplicative of them. You must hand in your paper as a written document (pen/pencil and paper). You also must hand in your printed-out readings.
When given this assignment, the first thing I thought about was functioning without my phone. When I am doing homework, I find myself constantly checking my phone. After reading "The Duggler's Brain" and "Your Brain is Evolving Right Now", I realized the control my phone has over me. The younger generation has a constant distraction because of the digital advancements made for brain stimulation, that they are unaware of. Through claims made from Carr and Vorgan, we can see that the internet has altered our actions, social skills, and routine.

I would like to start by discussing our actions. I began to notice that my addiction to my phone has gotten worse. I take my phone with me everywhere, even the bathroom. I also noticed that some of the things I do were in Carr's article. For example, Carr says "we tend to repeat the same actions over and over again." (2) Sometimes I find myself opening the same social application in a cycle. I will open Facebook, Snapchat, Twitter, and Instagram maybe ten to fifteen times a day.
“Physiology InsideOuts” in class kick off new units

Students are told to close their eyes and focus on the sound—no multitasking

Short (2-4 minute) versions at the start of class help students be present in class and see how that day’s topic fits into the larger picture of the body system we’re studying
A mind is a fire to be kindled, not a vessel to be filled.

--Plutarch
Great Books!