Exploring the Asymmetry of Metacognition

Ani Aghababyan
McGraw-Hill Education
281 Summer Street, 7th Floor
Boston, Massachusetts
USA
ani.aghababyan
@mheducation.com

Nicholas Lewkow
McGraw-Hill Education
281 Summer Street, 7th Floor
Boston, Massachusetts,
USA
nicholas.lewkow
@mheducation.com

Ryan Baker
University of Pennsylvania
3700 Walnut St.
Philadelphia, Pennsylvania
USA
rybaker@upenn.edu

ABSTRACT

People in general and students in particular have a tendency to misinterpret their own abilities. Some tend to underestimate their skills, while others tend to overestimate them. This paper investigates the degree to which metacognition is asymmetric in real-world learning and examines the change of a students' confidence over the course of a semester and its impact on the students' academic performance.

Our findings, conducted using 129,644 students learning in eight courses within the LearnSmart platform, indicate that poor or unrealistic metacognition is asymmetric. These students are biased in one direction: they are more likely to be overconfident than underconfident. Additionally, while the examination of the temporal aspects of confidence reveals no significant change throughout the semester, changes are more apparent in the first and the last few weeks of the course. More specifically, there is a sharp increase in underconfidence and a simultaneous decrease in realistic evaluation toward the end of the semester. Finally, both overconfidence and underconfidence seem to be correlated with students' overall course performance. An increase in overconfidence is related to higher overall performance, while an increase in underconfidence is associated with lower overall performance.

Keywords

confidence, metacognition, achievement, performance, discipline difference, learnign analytics, big data

1. INTRODUCTION

1.1 Background

Currently many educational digital environments customize a student's path to completion and mastery by allowing them to focus on content that they do not know and skip over parts that they believe they have already mastered. However, what if these assertions of knowledge are simply demonstrations of *overconfidence*? What if student's judgment of their abilities, their feeling of knowing [2, 9], is inaccurate? What if this overestimation of abilities creates an unrealistic expectation for the course, thus, discouraging the students and damaging their attitude toward their teachers or the pedagogy [15, 20]? If we want to customize based on students' perception of their abilities, we need to understand how their perception correlates to their actual knowledge and performance.

There is an extensive body of work that investigates and discusses students metacognitive experiences, particularly student confidence. Confidence has been defined as "the ability to believe in oneself" where this "belief" is considered to be learned [6]. Some researchers suggest a connection between students' confidence levels and their motivation, where "initial level of confidence and subsequent changes may affect one's motivation, performance, and possibly knowledge retention" [4]. However, many view confidence as-task specific metacognitive experience [7].

According to some studies, student academic success is dependent on many factors one of which is their confidence in being capable to succeed [3, 8, 10, 12, 13, 19]. In fact, most of these studies suggest that confidence is a reliable predictor of performance and success [18, 19, 21]. In his extensive meta-analyses, John Hattie identified student self-reported grades as being the factor most correlated with student achievement [11]. Hattie [11] suggests that if we manage to help the students outperform their own expectations, it could lead to higher grades. As confidence also reflects an expression of self-evaluation, the same could apply to confidence.

While research shows that confidence is a continuous concept and may range from low to high levels [17], two constructs in particular have emerged within this body of research as being important: overconfidence and underconfidence. These are the two ways that a student's estimation of their abilities can fail to be realistic. Current theories of student motivation suggest that if a student is overconfident, they may study less than if they possessed more accurate perceptions [16]. Perhaps overconfidence develops due to students' past positive grade experiences, which leads to their assumption that they will perform equally well in a new topic. As a result, they remain unaware of their need to adjust or develop their study skills [5]. By contrast, underconfidence may stem from a student's lack of self-assurance and belief in their own abilities. Findings suggest that encouraging realistic expectations and boosting academic confidence may benefit these students, leading to better performance

LAK '17, March 13-17, 2017, Vancouver, BC, Canada

DOI: http://dx.doi.org/10.1145/3027385.3027388

[14]. Hence, both overconfidence and underconfidence need to be monitored.

1.2 Study Goals

The purpose of this paper is to examine student confidence across eight courses. We aim to explore how symmetric confidence is (how balanced it is between overconfidence and underconfidence) in real-world learners. We will also study how these two manifestations of inaccurate confidence relate to student performance in the course, and how confidence evolves throughout the semester. In addition, we will examine whether student metacognition is differentially successful in different subject domains.

2. METHODS

2.1 Materials

This study was conducted using data collected from one of McGraw-Hill Education's learning platforms, LearnSmart. LearnSmart is an adaptive learning program that personalizes learning and provides study paths for students. Within this environment students access their course materials, learn and practice the content, and complete assignments. Over 5.9 billion questions have been answered since 2009.

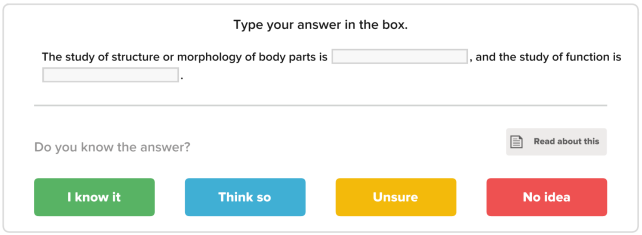

Figure 1: LearnSmart Interface

For this study, we selected data from eight different courses that were taught in the Spring 2015 semester using the platform. Our selection included four courses from humanities/social sciences and four from the physical/life sciences. To ensure that these courses were comparable in the number of total questions answered throughout the semester, we selected eight of the courses that had relatively equivalent usage in 2015: Spanish, Psychology, Introduction to Business, Management, Practical Introduction of Medical Assisting, Anatomy & Physiology, Biology, and General Chemistry.

2.2 Participants

Due to the regulations regarding student data collection and usage, our platform does not collect gender, ethnicity, or other demographic information from our participants. Every student who submitted at least 1 assignment in their course were included in this study, for a total of N=130,791 students and 102,082,551 item responses by students.

2.3 Measures of Confidence

In its efforts to improve educational outcomes, the field of education faces obstacles such as the abundance of multiple-choice tests that reinforce students' guesswork behavior and in the meantime fail to measure the degree of confidence that students exhibit towards their knowledge [1]. Relatively few learning systems measure the confidence that students exhibit towards their knowledge [1]. In an attempt to implement confidence measurement, LearnSmart asks students their perception of their confidence alongside each content question, avoiding delay in response or recall bias. (see Figure 1).

For each question within each assignment, students are asked to provide an answer. Using the confidence scale embedded within the interface, before submitting each answer students were prompted to report their confidence level on a four-item scale: "I know it" (64.7% of the data), "Think so" (27.7% of the data), "Unsure" (5.6% of the data), "No Idea" (5.5% of the data).

2.4 Measures of Accuracy

In addition to the confidence metric, the interface also recorded the correctness (otherwise known as the score) of the student's answer. Each student response was automatically graded according to the following categories: incorrect, partially correct (this is only for items with dual questions, which comprise approximately 5% of total questions), and correct. These responses were given three possible scores; 0 (incorrect, which was 32.8% of the data), 1 (partially correct, which was 5.3% of the data), and 2 (correct, which was 65.3% of the data).

Due to lack of a final score metric within the database, we calculated our own total grade for each student. We calculated each student's accuracy score by dividing the number of correct answers by the total number of questions answered by the student. The result showed a mean score of 69%.

3. ANALYSES & RESULTS

3.1 Data Exploration

Prior to exploring the data for confidence profiles, we removed responses where the student had received partial credit (about 5% of all questions), as these responses were ambiguous for the analyses we will present below. Additionally, we excluded rows with "think so" and "unsure" reports of confidence, as these responses were not indicative of overconfidence or underconfidence. As a result, our total number of items was reduced to 68,363,910 with a total of N=129,644 unique users, and a total of 51,657 unique questions answered. The average number of questions answered per user was around 424.

3.2 Confidence Profiles

To begin our analyses on confidence, we operationalized students' confidence profiles as seen in Figure 2. From this diagram, we calculated overconfidence and underconfidence respectively by calculating the conditional probabilities of student being confident (confidence = 3) when their answer was incorrect (score = 0) and of student being not confident (confidence = 0), when their answer was correct (score = 2). As discussed below, we also analyze this separately for the different categories of courses: humanities/social or physical/life science.

To identify the proportions of overconfidence, underconfidence, realistic, and knowledgeable beliefs, we used the general conditional probability formula as seen below:

$$\textbf{Realistic} = P(c = 0 \ \& \ s = 0)/n$$
$$\textbf{Underconfidence} = P(c = 0 \ \& \ s = 2)/n$$

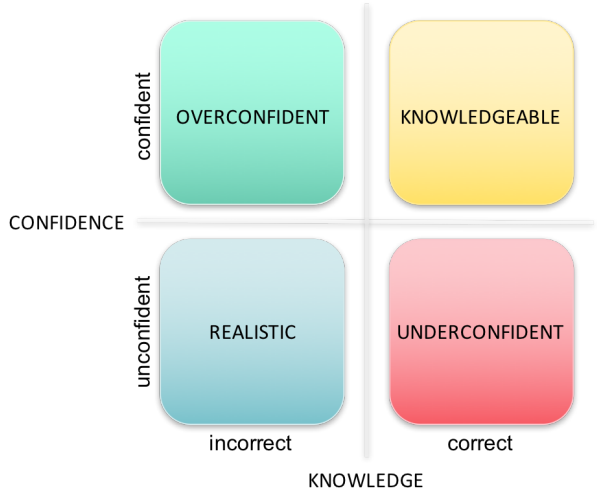

Figure 2: Confidence Profiles

Overconfidence= P(c = 3 & s = 0)/n
Knowledgeable= P(c = 3 & s = 2)/n

where **c** is confidence, **s** is score, and **n** is the number of questions in our sample.

We then create three tables to see the overall prevalence of each category: confidence profiles for all courses combined, confidence profiles for courses in physical sciences, and confidence profiles for courses in humanities/social sciences (see Tables 1, 2, and 3).

Table 1: Confidence profile for all courses

Confidence Profile	Count	Percent
Realistic	5,469,675	8%
Under-confident	161,759	0.24%
Over-confident	15,526,737	**22.71%**
Knowledgeable	47,205,739	69.05%

Table 2: Confidence profile for physical/life science courses

Confidence Profile	Count	Percent
Realistic	3,262,840	9.88%
Under-confident	53,861	0.16%
Over-confident	8,496,751	**25.73%**
Knowledgeable	21,209,445	64.23%

Table 1 shows that 22.71% of the time students are overconfident in their abilities; by contrast, they are only underconfident 0.24% of the time. This suggests that students are more likely to overestimate their abilities than to underestimate their abilities. The same pattern repeats itself in both groups of courses (25.73% vs 0.16% in physical sciences and 19.89% vs. 0.31% in humanities/social sciences), which suggests that despite the difference in discipline, students

Table 3: Confidence profile for humanities/social science courses

Confidence Profile	Count	Percent
Realistic	2,206,835	6.24%
Under-confident	107,898	0.31%
Over-confident	7,029,986	**19.89%**
Knowledgeable	25,996,294	73.56%

metacognition is indeed asymmetric within LearnSmart and it leans toward overestimation of abilities. However, there is a difference in how much overconfidence is seen by discipline. 25.73% of the time students are overconfident in their abilities in physical/life sciences vs. 19.89% of the time shown in humanities/social sciences. This difference in proportions was very large; students in the physical/life sciences were overconfident 29.3% more often (we do not present a statistical test due to the massive data set size; virtually any difference would be statistically significant).

3.3 Temporal Representation of Confidence

We can understand how confidence changes over the course of the semester, by visualizing the proportion of each combination of accuracy and confidence, shown in Figure 3. In this Figure, the x axis is the weeks of the semester starting from January 1st and ending on June 15th, while the y axis is the log scale of the percentage of questions answered which had each category of reported confidence in that week.

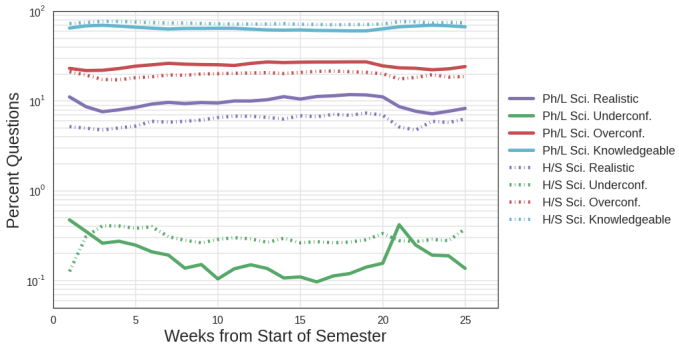

Figure 3: Temporal Representation of Confidence for Physical/LIfe and Humanities/Social Sciences

In this figure we can notice several interesting changes. The underconfidence for both discipline types is low throughout the semester. However, at the beginning of the semester, between weeks 2 and 3, when underconfidence rate rises in humanities/social sciences, the underconfidence in physical sciences drops. A similar shift happens at the end of the semester, toward week 21, when the underconfidence rate in physical sciences rises considerably while the rate drops for the humanities/social sciences. Additionally, in the same week (week 21), when underconfidence rate in physical sciences rises dramatically, the rate of students being realistic drops both for physical/life and humanities/social sciences. It is possible that this indicates that at this point students are more worried about their final grade and the sufficiency of their preparedness.

3.4 Correlations: Accuracy score vs. Confidence Profiles

Finally, we explore the relationship between students' overall course performance and their reported confidence levels. For this purpose, we calculated the proportion of correctly answered questions that were underconfident for each student. Similarly, the overconfidence ratio was calculated as the proportion of incorrectly answered questions that were overconfident per student. Then we correlated these two new variables with the students' overall course performance. Pearson correlation coefficient revealed that higher overconfidence seems to be correlated to higher scores, while higher underconfidence is negatively correlated with success. In addition, this correlation is larger in magnitude for humanities/social sciences (see Tables 5 and 2). We also calculated the Spearman correlation coefficients, which were lower across all categories. This emphasizes a larger linear correlation as opposed to a rank correlation. It can be explained by the skewed nature of our data and the influential observations in the tails of the distribution.

Table 4: Accuracy vs. Confidence Correlation Results for All Courses

Confidence Profile	Pearson Correlation
Underconfidence Ratio vs. Accuracy	-0.227
Overconfidence Ratio vs. Accuracy	0.489

Table 5: Accuracy vs. Confidence Correlation Results for Physical Science Courses

Confidence Profile	Pearson Correlation
Underconfidence Ratio vs. Accuracy	-0.189
Overconfidence Ratio vs. Accuracy	0.445

Table 6: Accuracy vs. Confidence Correlation Results for Humanities/Social Science Courses

Confidence Profile	Pearson Correlation
Underconfidence Ratio vs. Accuracy	-0.279
Overconfidence Ratio vs. Accuracy	0.516

We chose scatter plots to display the correlation results for all courses using a color legend to visualize the course categories within each plot. Figure 4 demonstrates a medium positive correlation between students' overconfidence ratio and their scores. The opposite is visible in Figure 5 where we see a small negative correlation between students' overconfidence ratio and their scores. For underconfidence ratio plot, we retained only the data from students that had at least one question where they demonstrated underconfidence. In addition to the general correlation pattern, these plots also reveal a pattern where students in physical/life science courses are consistently more accurate in their estimations of their ability than students in humanities/social science courses).

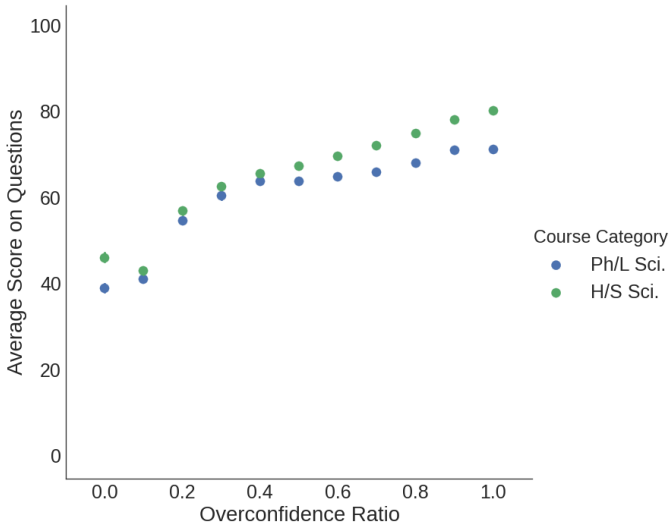

Figure 4: Correlation of Accuracy vs. Overconfidence Ratio For Physical/LIfe and Humanities/Social Science Courses

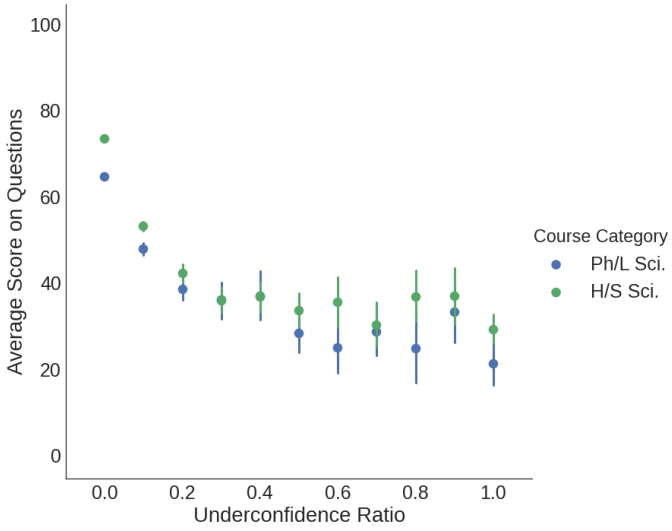

Figure 5: Correlation of Accuracy vs. Underconfidence Ratio For Physical/LIfe and Humanities/Social Science Courses

4. DISCUSSION & FUTURE WORK

A large number of research studies have already asserted the importance of students' metacognition and confidence. Hence, learning how students' confidence interacts with their performance, how it evolves throughout the course, and how it varies from discipline to discipline can bring important insights to monitoring and helping students learn to regulate it.

In this paper we have explored students' academic confidence. We created four confidence profiles and discovered that students' perception of their abilities in real-world learning is asymmetric; students are much more likely to be overconfident than underconfident. This pattern is even more pronounced for physical/life sciences than for other

courses. We also explored the change of confidence over the course of the semester, noting increased variability in the levels of confidence at the beginning and at the end of the semester. Finally, the results from this work support previous findings that students' perception of their performance is in fact correlated with their actual performance. However, we find that overconfident students perform relatively well. This finding suggests that some of this seeming overconfidence may actually represent slips; the student may really have known the skill despite getting the answer wrong. Alternatively, estimations of skill may be general rather than pertaining to the current situation. By contrast, students who were underconfident generally did worse. Whether this implies that underconfident students should become more confident, or that they need more help, is a relevant area for future research. We also found that both forms of inaccurate confidence are more prominent in humanities/social science courses than in physical/life science courses. It is possible that this is because it is easier to estimate one's proficiency on procedural skills than on factual matters; this is also a relevant area for future work.

There are several other future directions that will also be valuable for expanding scientific understanding of these phenomena. First, it may be worth incorporating measures of item difficulty into these analyses to see how it influences over/underconfidence. In addition, it would be valuable to increase the number of courses within each discipline category for more rigorous investigations into how discipline impacts over/underconfidence; similarly, breaking down different types of material will help us to explore whether discipline-level effects are due to disciplnary culture or due to the types of material being studied. Finally, these results suggest that it may be worth developing interventions to help students be more realistic about what they do not know, within platforms such as LearnSmart.

5. ACKNOWLEDGMENTS

This paper is based on work supported by McGraw-Hill Education. We would like to extend our appreciation to our CDO Stephen Laster, the analytics VP Alfred Essa, and to the Learn Smart team for the research support provided to complete this project. Any opinions, findings, or recommendations expressed in this paper are those of the authors and do not necessarily reflect positions or policies of the company.

6. REFERENCES

[1] T. M. Adams and G. W. Ewen. The importance of confidence in improving educational outcomes. In *25th Annual Conference on Distance Teaching & Learning, Madison, WI.*, 2009.

[2] R. Azevedo and J. Cromley. Does training on self-regulated learning facilitate students' learning with hypermedia? *Journal of educational psychology.*, 96(3):523, 2004.

[3] A. Bandura. Self-efficacy: toward a unifying theory of behavioral change. *Psychological Review.*, 84(2):191, 1977.

[4] M. Besterfield-Sacre, N. Y. Amaya, L. J. Shuman, C. J. Atman, and R. L. Porter. Understanding student confidence as it relates to first year achievement. In *Frontiers in Education Conference, 1998. FIE'98. 28th Annual*, volume 1, pages 258–263. IEEE, 1998.

[5] D. E. Clayson. Performance overconfidence: metacognitive effects or misplaced student expectations? *Journal of Marketing Education.*, 27(2):122–129, 2005.

[6] L. Copeland. Developing student confidence: the post clinical conference. *Nurse Educator.*, 15(1):7, 1990.

[7] A. Efklides. Interactions of metacognition with motivation and affect in self-regulated learning: The masrl model. *Educational psychologist.*, 46(1):6–25, 2011.

[8] K. A. Ericsson, R. T. Krampe, and C. Tesch-Römer. The role of deliberate practice in the acquisition of expert performance. *Psychological review.*., 100(3):363, 1993.

[9] J. Hart. Memory and the feeling-of-knowing experience. *Journal of educational psychology.*, 56(4):208, 1965.

[10] S. Harter. Effectance motivation reconsidered. toward a developmental model. *Human development.*, 21(1):34–64, 1978.

[11] J. Hattie. *Visible learning: A synthesis of over 800 meta-analyses relating to achievement.* Routledge, 2008.

[12] J. Kuhl. Volitional aspects of achievement motivation and learned helplessness: Toward a comprehensive theory of action control. *Progress in experimental personality research.*, 13:99–171, 1984.

[13] J. Nicholls. Achievement motivation: Conceptions of ability, subjective experience, task choice, and performance. *Psychological review.*, 91(3):328, 1984.

[14] L. Nicholson, D. Putwain, L. Connors, and P. Hornby-Atkinson. The key to successful achievement as an undergraduate student: confidence and realistic expectations? *Studies in Higher Education.*, 38(2):285–298, 2013.

[15] M. Nicolaidou and G. Philippou. Attitudes towards mathematics, self-efficacy and achievement in problem solving. *European Research in Mathematics Education III. Pisa: University of Pisa.*, 2003.

[16] C. Nowell and R. M. Alston. I thought i got an a! overconfidence across the economics curriculum. *The Journal of Economic Education.*, 38(2):131–142, 2007.

[17] B. D. Pulford. *Overconfidence in human judgement.* University of Leicester, 1986.

[18] L. Stankov, J. Lee, W. Luo, and D. J. Hogan. Confidence: A better predictor of academic achievement than self-efficacy, self-concept and anxiety? *Learning and Individual Differences.*, 22(6):747–758, 2012.

[19] R. Stiggins. Assessment, student confidence, and school success. *The Phi Delta Kappan.*, 81(3):191–198, 1999.

[20] T. A. Stinson and X. Zhao. Unmet expectations: why is there such a difference between student expectations and classroom performance? *Journal of College Teaching & Learning (TLC).*, 5(7), 2011.

[21] A. Zajacova, S. M. Lynch, and T. J. Espenshade. Self-efficacy, stress, and academic success in college. *Research in higher education.*, 46(6):677–706, 2005.