

Two-Stage Collaborative Exams

Iris Howley

www.irishowley.com

Postdoctoral Research Fellow

Graduate School of Education

Lytics Lab @ Stanford

“If you ask someone else for help on a problem in an exam, you are cheating, but if you don’t ask someone for help on a problem in the real world, you are a fool.” – Dan Schwartz

* I’m not an expert, just interested in the topic of Two-Stage Collaborative Exams!

Required Readings

1. Gilley & Clarkston (2014). "Collaborative testing: Evidence of learning in a controlled in-class study of undergraduate students." <http://www.jstor.org/stable/43632038>
2. 3 Examples of Two Stage Exams in Physics Classrooms: (1) Rieger & Heiner (2014). "Examinations that support collaborative learning: the students' perspective." <http://www.jstor.org/stable/43632011> (2) Weiman et al. (2014). "Physics exams that promote collaborative learning." <http://aapt.scitation.org/doi/10.1119/1.4849159> (3) Jang et al. (2017). "Collaborative exams: Cheating? Or learning?" <http://aapt.scitation.org/doi/10.1119/1.4974744>
3. Smith et al. (2009) "Why peer discussion improves student performance on in-class concept questions." <http://science.sciencemag.org/content/323/5910/122>

Supplemental:

- (psychology) Zimbardo et al (2003). Cooperative college examinations: More gain, less pain when students share information and grades. <http://www.tandfonline.com/doi/abs/10.1080/00220970309602059>
- (sociology) Zipp (2007). Learning by exams: The impact of two-stage cooperative tests. <http://journals.sagepub.com/doi/abs/10.1177/0092055X0703500105>
- (nursing) Lusk & Conklin (2003). Collaborative testing to promote learning. <http://search.proquest.com/docview/203944535>
- (biology) Leight et al (2012). Collaborative testing improves performance but not content retention in a large-enrollment introductory biology class. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3516795/>

Required Readings

1. Gilley & Clarkston (2014). “Collaborative testing: Evidence of learning in a controlled in-class study of undergraduate students.”

Experimental foundation, Controlled, Benefits

2. (1) Rieger & Heiner (2014). “Examinations that support collaborative learning: the students' perspective.” (2) Weiman et al. (2014). “Physics exams that promote collaborative learning.” (3) Jang et al. (2017). “Collaborative exams: Cheating? Or learning?”

Practical concerns, Logistics, Benefits

3. Smith et al. (2009) “Why peer discussion improves student performance on in-class concept questions.”

Why 2-Stage exams might increase understanding

Handouts

1. 1-3 Collaborative Exam Questions

- From Rieger & Heiner, Weiman et al., Jang et al

2. 5E Model of Instruction

- Not part of topic, but in use today!
- Approach for structuring a lesson plan

Some content adapted from Eric Mazur's "Assessment: The Silent Killer of Learning"

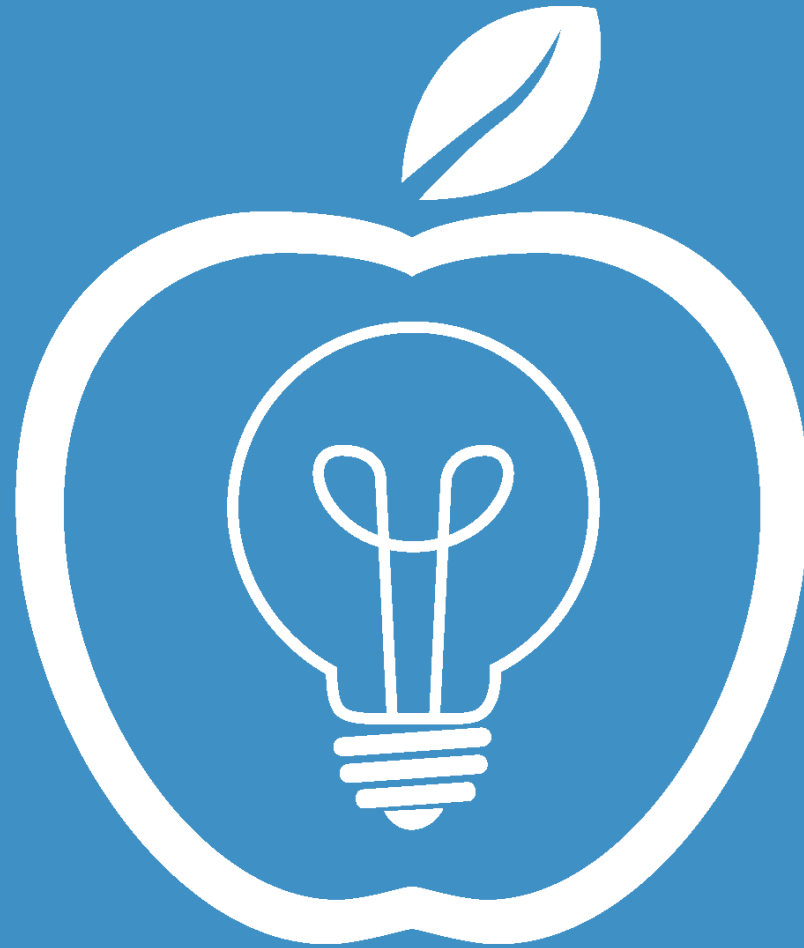
Learning Goals

By the end of our session, participants should be able to...

1. List several purposes of assessment
2. Summarize pros and cons of two-stage collaborative assessments
3. Describe logistics of two-stage collaborative exams



Backward
Design



Engage

generate interest, access prior knowledge, connect to prior knowledge, frame the idea

What is the purpose of assessment?

Individually. List as many purposes as you can.
(2 minutes)

With a partner. Discuss assessment purposes.
(3 minutes)

Purposes of Assessment

- Measure student understanding
- Self-assessment
- Teaching effectiveness
- Assessment effectiveness
- Give a grade/summative
- Accountability to study
- Certification
- Ranking
- Individualized feedback
- **LEARNING**





Explore

experience key concepts, discover new skills, probe/inquire/question experiences, examine thinking, establish relationships & understanding

Purposes of Assessment



Purposes of Assessment

<https://youtu.be/6sY7UJCYbNg> **(58:46)**



Explain

connect prior knowledge & background to new discoveries, communicate new understandings

What is a two-stage collaborative exam?

Two-Stage Collab Exams: Logistics

1. Gilley & Clarkston (1.1) More experimental, less practical
 - a. Individual: 45 minutes, 45 problems
 - b. Collab: 45 minutes, 40 problems
 - 3-5 students, student-selected
2. Reiger & Heiner + Weiman (2.1 & 2.2)
 - a. Individual: 3/4 – 2/3 exam time (75-90% grade)
 - b. Collab: 1/4 – 1/3 exam time (25-10% grade)
 - 3-4 students, preformed OR student-selected
 - Mostly same problems as Individual phase (extensions, isomorphic)
3. Jang, Lasry, Miller & Mazur (2.3)
 - a. Final exam score is mean of Individual + Collab phases
 - b. Correct answer: 4pts → 2pts → 1pt
 - c. 7-11 problems, no one scores above 50% on Individual phase
 - d. Mostly same problems on Individual + Collab phases

90 minute exam period seems standard, but possible in 60 minutes

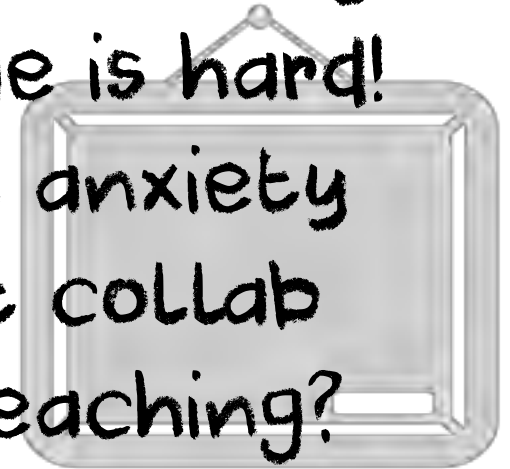
Two-Stage Collaborative Exams

Pros

- Helps bottom half of students
- Helps the top half
- Social performance
- Better at collab
- More depth

Cons

- Personal conflicts
- Free rider
- Fairness
- Accountability
- Change is hard!
- Social anxiety
- Bad at collab
- Not teaching?



Two-Stage Collaborative Exams

Pros

- Immediate feedback
- + Student learning
- + Student motivation
- + Retention of content?
- - Anxiety
- + Collaborative Skills
- + Course perceptions
- + Appreciate collabor.
- + Retention

Cons

- - Number test ques.
- + Administrative effort
- Individual accountable
 - Free-riders
 - Domineers
 - Propagate answers
- Resistant to change!

Why do two-stage collaborative exams increase student learning?

Smith, Wood, Adams, Wieman, Knight, Guild, & Su (2009). "Why Peer Discussion Improves Student Performance on In-Class Concept Questions." *Science*.

“Our results suggest that peer discussion can be effective for understanding difficult concepts even when no one in the group initially knows the correct answer.”

Smith, Wood, Adams, Wieman, Knight, Guild, & Su (2009). “Why Peer Discussion Improves Student Performance on In-Class Concept Questions.” *Science*.



Extend/Elaborate

apply new learning to a new situation, explain concept being explored, communicate new understanding with formal language

What concerns do you have about implementing two-stage collaborative exams in your classroom?

Individually. Spend 2 minutes.

Pair. Discuss with a partner. 3 minutes.



Evaluate

assess understanding (self, peer and teach evaluation), demonstrate understanding of new concept by observation or open response, apply within problem situation, show evidence of accomplishment

Office of Postdoctoral
Affairs will send you an
online evaluation
survey to complete

(It'd be super great if you could fill that out!)



Iris Howley
www.irishowley.com
Postdoctoral Research Fellow
Lytics Lab @ Stanford



dents (i.e., a student's performance relative to the class), yet even the small weight of the group portion provides sufficient motivation for students to take this part seriously. For example, an 85/15 (individual/group) split used in our physics class

resulted in an average increase of the midterm mark due to the group portion of 3.3% and an average increase in the final exam score due to the group portion of 1.6%. The resulting impact on the average course grade of the group part of the exams was

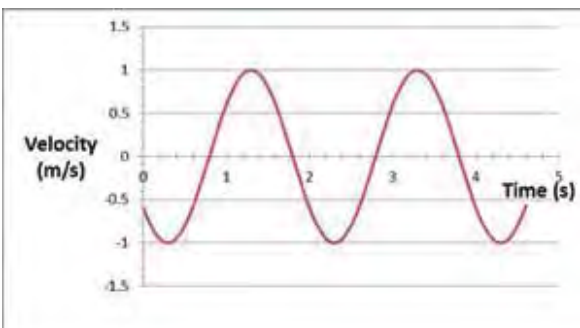
0.5% from the midterm and 0.7% for the final exam, where the standard deviation of course grade distribution was 9.7%.

On the basis of the collective experience at UBC across the science disciplines of physics, chemistry,

FIGURE 1

Examples of questions taken from a two-stage exam for physics.

Most questions will be the same for the individual and the group part. If questions are modified, it is usually to reduce the number of detailed calculations, which do not promote discussions, and replace with prompts to "explain your reasoning." Additionally, one or two more challenging questions may be added.

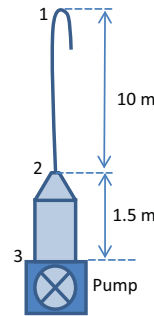
INDIVIDUAL PART	GROUP PART
<p>A train is approaching the train station at velocity $v_0 = 15$ m/s relative to the ground in still air. The train operator sounds the train whistle, which emits a note with frequency $f_0 = 2500$ Hz.</p> <p>The sound of the whistle is heard by different observers:</p> <p>The train operator hears a frequency f_A;</p> <p>a person standing on the station platform watching the train approach hears a frequency f_B;</p> <p>the operator of a second train approaching the station from the other direction with velocity $v_2 = 10$ m/s hears a frequency f_C.</p> <p>What are the frequencies f_A, f_B, and f_C?</p>	<p>(Changed to ranking) A train is approaching the train station at velocity v_0 relative to the ground in still air. The train operator sounds the train whistle, which emits a note with frequency f_0.</p> <p>The sound of the whistle is heard by different observers:</p> <p>The train operator hears a frequency f_A;</p> <p>a person standing on the station platform watching the train approach hears a frequency f_B;</p> <p>the operator of a second train approaching the station from the other direction with velocity v_2 hears a frequency f_C;</p> <p>a passenger traveling on a slower train that has just been overtaken by the first train (and sees the first train move farther away) hears frequency f_D.</p> <p>Rank the frequencies heard by the observers (f_A, f_B, f_C, f_D) in order from the highest to the lowest frequency.</p>
<p>The graph shows the velocity vs. time graph of a harmonic oscillator.</p>  <p>Determine</p> <ol style="list-style-type: none"> the angular frequency the maximum displacement the phase constant and the equation describing the position as a function of time. 	<p>(Replace part c)</p> <ol style="list-style-type: none"> same same Sketch the potential energy curve as a function of time. Assume that we have a horizontal harmonic oscillator.

Box 1. Examples of questions taken from a two-stage exam for physics.

Most questions will be the same for the individual and the group part. If questions are modified, it is usually to reduce the number of detailed calculations, which do not promote discussions, and replace with prompts to “explain your reasoning.” Additionally, one or two more challenging questions may be added.

Question

Assume you want to design a water fountain for your local park. The fountain is supposed to shoot water up to a height of 10.0 m above the exit nozzle, which is located 1.5 m above a pump that pumps water into a vertical tube of 5.0 cm diameter. The pump has a gauge pressure of 100 kPa.

**Individual Part**

- Rank the pressures at points 1 (at the top), 2 (at the exit of the nozzle), and 3 (at the exit of the pump).
- What is the diameter of the exit nozzle?

Group Part

- Part b changed to ranking:
- Rank the velocities at points 1, 2, and 3.

Question

You and your little sister are out in the snow on a sled that has a mass of 11 kg. Your sister, who weighs 29 kg, is sitting on the sled and you want to push her along. You start applying a horizontal force and initially the sled doesn't move but you slowly increase your force until, suddenly, the sled does move. You maintain the same force that you were applying when the sled started moving for the next 5.0 s after which you let go. (Assume that the kinetic friction coefficient is $\mu_k = 0.02$ and the static friction coefficient is $\mu_s = 0.08$ in this case.)

Individual Part

- How far do you have to run if you apply the force for 5.0 s?
- What is your sister's speed at $t = 5.0$ s?
- After letting go, how far do your sister and her sled move until she is stationary again?
(In case you could not solve part b, assume that her speed is $v = 2.5$ m/s at $t = 5.0$ s.)

Group Part

- (Converting calculation to reasoning and representation with graphs.)
- Draw a qualitative diagram that roughly shows the net force acting on the sled as a function of time. (Qualitative means that it explains the overall behavior without using exact numbers.)
 - Draw a second qualitative graph of the acceleration of the sled as a function of time.
 - Draw a third qualitative graph of the velocity of the sled as a function of time.

the two-stage format. Students also see the benefits of these discussions. We rarely have to discourage students from working individually during the group portion, and students that are usually too shy to speak up during in-class activities will defend their answers vigorously during the second stage of the exam. As confirmed through both observations and student self-reports,⁵ a large fraction of the groups discuss the questions until all members agree on an answer, or they take a vote in cases where an agreement cannot be achieved. The high stakes context of an exam combined with the fact that all students are well prepared to participate in the discussion, because (a) they have studied for the exam and (b) they thought carefully about the questions and committed to an answer just moments ago during the individual portion, produce the perfect environment for rich discussion. Although we introduce collaborative learning activities into the course before the exams and explain the benefits, for many students

the value becomes more readily apparent during the two-stage exam.

We see this on survey responses and in the behavior of the class after the first two-stage exam. Students' response to the use of two-stage exams is overwhelmingly positive, with 87% of the students recommending continued use of two-stage midterm exams and only a few percent recommending against their use. Examples of typical positive comments are:

Student A: “I was able to instantly learn from my mistakes.”

Student B: “It was good to compare methods and answers with others, and it allowed us to be more confident.”

Student C: “Interesting. All had different ways [of] approaching the question. Very helpful to understand everyone's response and why they thought their answer was correct.”

platform that automatically grades open-ended and multiple-choice questions and manages team assessments by assigning groups and providing iterative feedback. Given that instructors need only input assessment questions and that the management and grading are fully automated, this system makes the effort involved in implementing collaborative assessments comparable to that of conventional exams. While our implementation of collaborative assessments might involve more exam time over the course of the semester than in traditional physics courses (90 minutes per assessment, 5 times during the semester), the pedagogical purpose of these exams is not simply to assess students but also to provide students with an opportunity to discuss problems with one another and learn through this experience.

Given these findings, it is up to the readers of this paper to help overcome the fourth barrier: resistance to changing established practices. Although we do not address the nature or the complexity of the knowledge and skills acquired during collaborative exams, we show that even in institutions where established practices have a very long history, collaborative exams can be effectively implemented with significant benefit to all students. We hope the realization that collaboration can turn assessment into a learning opportunity will encourage instructors to adopt collaborative assessment practices more broadly.

ACKNOWLEDGMENTS

Several people contributed to this work. H.J. conceived of the basic idea for this work. E.M., K.M., and H.J. designed and carried out the experiments, and analyzed the results. E.M. supervised the research and the development of the manuscript. H.J. and N.L. co-wrote the manuscript; all authors subsequently took part in the revision process and approved the final copy of the manuscript. E.M. and K.M. provided feedback on the manuscript throughout its development. This research was supported by the National Science Foundation under Grant No. NSF DUE 1504664.

APPENDIX: SAMPLE ASSESSMENT QUESTIONS

1. Multiple-choice question

A classmate leaves a message on your voice mail betting that you cannot throw a stone hard enough so it lands on the roof of a 20-m high building. As you stare out of your window pondering whether to accept the challenge, the well in the courtyard suddenly gives you an idea. You drop a stone into the well and note that you hear a splash 4.0 s later. You repeat the experiment with another stone, but this time, you throw the stone down as fast as you can. This time the splash comes 3.0 s after the stone leaves your hand. Armed with this information you carry out a quick calculation and then you call back your friend. Do you accept the bet?

- (A) Yes, but it is close
- (B) Yes, easily
- (C) No, but it is close
- (D) No, not by a wide margin
- (E) Insufficient information in this problem

2. Open-ended question

Three books, each of inertia m , rest on the floor of an elevator. The elevator starts at the first floor and rises to the sixth floor. It travels at a constant speed between the second and fifth floors, as it rises by a total distance h . Enter an expression for the work done by the bottom book on the middle book during the passage from the second to the fifth floors in terms of m , h , and the acceleration due to gravity g .

^{a)}Electronic mail: wanypie@gmail.com

^{b)}Electronic mail: nlasry@me.com

^{c)}Electronic mail: kellymillervt@gmail.com

^{d)}Electronic mail: eric.mazur@gmail.com

¹P. Heller, R. Keith, and S. Anderson, "Teaching problem solving through cooperative grouping. Part 1: Group versus individual problem solving," *Am. J. Phys.* **60**, 627–636 (1992).

²P. Heller and M. Hollabaugh, "Teaching problem solving through cooperative grouping. Part 2: Designing problems and structuring groups," *Am. J. Phys.* **60**, 637–644 (1992).

³L. S. Vygotsky and Michael Cole, *Mind in Society: The Development of Higher Psychological Processes* (Harvard U.P., Cambridge, MA, 1978).

⁴W. Damon and E. Phelps, "Critical distinctions among three approaches to peer education," *Int. J. Educ. Res.* **13**, 9–19 (1989).

⁵G. W. Rieger and C. E. Heiner, "Examinations that support collaborative learning: The students' perspective," *J. Coll. Sci. Teach.* **43**, 41–47 (2014).

⁶H. Shirouzu, N. Miyake, and H. Masukawa, "Cognitively active externalization for situated reflection," *Cognit. Sci.* **26**, 469–501 (2002).

⁷P. Black and D. Wiliam, "Assessment and classroom learning," *Assess. Educ.* **5**, 7–74 (1990).

⁸J. P. Murry, "Better testing for better learning," *Coll. Teach.* **38**, 148–152 (1998).

⁹B. Gilley and B. Clarkston, "Collaborative testing: Evidence of learning in a controlled in-class study of undergraduate students," *J. Coll. Sci. Teach.* **43**, 83–91 (2014).

¹⁰M. Lusk and L. Conklin, "Collaborative testing to promote learning," *J. Nurs. Educ.* **42**, 121–124 (2003).

¹¹M. K. Smith *et al.*, "Why peer discussion improves student performance on in-class concept questions," *Science* **323**, 122–124 (2009).

¹²C. E. Wieman, G. W. Rieger, and C. E. Heiner, "Physics exams that promote collaborative learning," *Phys. Teach.* **52**, 51–53 (2014).

¹³J. G. Lambiotte *et al.*, "Cooperative learning and test taking: Transfer of skills," *Contemp. Educ. Psychol.* **12**, 52–61 (1987).

¹⁴J. V. Shindler, "Greater than the sum of the parts? Examining the soundness of collaborative exams in Teacher Education Courses," *Innovative Higher Educ.* **28**, 273–283 (2004).

¹⁵P. G. Zimbardo, L. D. Butler, and V. A. Wolfe, "Cooperative college examinations: More gain, less pain when students share information and grades," *J. Exp. Educ.* **71**, 101–125 (2003).

¹⁶S. Sandahl, "Collaborative testing as a learning strategy in nursing education," *Nurs. Edu. Perspect.* **31**, 142–147 (2010).

¹⁷S. A. Stearns, "Collaborative exams as learning tools," *Coll. Teach.* **44**, 142–147 (2010).

¹⁸"Learning Catalytics Home Page," <<https://learningcatalytics.com>>.

¹⁹L. K. Michaelsen, M. Sweet, and D. X. Parmelee, *Team-Based learning: Small Group Learning's Next Big Step: New Directions for Teaching and Learning* (Wiley, Lexington, KY, 2011), Vol. 116.

²⁰P. C. Blumenfeld *et al.*, "Motivating project-based learning: Sustaining the doing, supporting the learning," *Educ. Psychol.* **26**, 369–398 (1991).

²¹C. H. Crouch and E. Mazur, "Peer instruction: Ten years of experience and results," *Am. J. Phys.* **69**, 970–977 (2001).

²²D. Hestenes, M. Wells, and G. Swackhamer, "Force concept inventory," *Phys. Teach.* **30**, 141–158 (1992).

²³J. F. Zipp, "Learning by exams: The impact of two-stage cooperative tests," *Teach. Sociol.* **35**, 62–76 (2007).

The 5E Model of Instruction



5E Definition	Teacher Behavior	Student Behavior
Engage		
<ul style="list-style-type: none"> • Generate interest • Access prior knowledge • Connect to past knowledge • Set parameters of the focus • Frame the idea 	<ul style="list-style-type: none"> • Motivates • Creates interest • Taps into what students know or think about the topic • Raises questions and encourages responses 	<ul style="list-style-type: none"> • Attentive in listening • Ask questions • Demonstrates interest in the lesson • Responds to questions demonstrating their own entry point of understanding
Explore		
<ul style="list-style-type: none"> • Experience key concepts • Discover new skills • Probe, inquire, and question experiences • Examine their thinking • Establish relationships and understanding 	<ul style="list-style-type: none"> • Acts as a facilitator • Observes and listens to students as they interact • Asks good inquiry-oriented questions • Provides time for students to think and to reflect • Encourages cooperative learning 	<ul style="list-style-type: none"> • Conducts activities, predicts, and forms hypotheses or makes generalizations • Becomes a good listener • Shares ideas and suspends judgment • Records observations and/or generalizations • Discusses tentative alternatives
Explain		
<ul style="list-style-type: none"> • Connect prior knowledge and background to new discoveries • Communicate new understandings • Connect informal language to formal language 	<ul style="list-style-type: none"> • Encourages students to explain their observations and findings in their own words • Provides definitions, new words, and explanations • Listens and builds upon discussion form students • Asks for clarification and justification • Accepts all reasonable responses 	<ul style="list-style-type: none"> • Explains, listens, defines, and questions • Uses previous observations and findings • Provides reasonable responses to questions • Interacts in a positive, supportive manner
Extend/Elaborate		
<ul style="list-style-type: none"> • Apply new learning to a new or similar situation • Extend and explain concept being explored • Communicate new understanding with formal language 	<ul style="list-style-type: none"> • Uses previously learned information as a vehicle to enhance additional learning • Encourages students to apply or extend the new concepts and skills • Encourages students to use terms and definitions previously acquired 	<ul style="list-style-type: none"> • Applies new terms and definitions • Uses previous information to probe, ask questions, and make reasonable judgments • Provides reasonable conclusions and solutions • Records observations, explanations, and solutions
Evaluate		
<ul style="list-style-type: none"> • Assess understanding (Self, peer and teacher evaluation) • Demonstrate understanding of new concept by observation or open-ended response • Apply within problem situation • Show evidence of accomplishment 	<ul style="list-style-type: none"> • Observes student behaviors as they explore and apply new concepts and skills • Assesses students' knowledge and skills • Encourages students to assess their own learning • Asks open-ended questions 	<ul style="list-style-type: none"> • Demonstrates an understanding or knowledge of concepts and skills • Evaluates his/her own progress • Answers open-ended questions • Provides reasonable responses and explanations to events or phenomena