

References

Brown, H. D. & Abeywickrama, P. (2010). *Language assessment principles and classroom practices*. White Plains, NY: Pearson Education.

Food diary template. (n.d.). *Vertex42*. Retrieved from <https://www.vertex42.com/ExcelTemplates/food-diary-template.html>

Paulsen, R. (n.d.). Civil rights progress: 1954-1958. *Pinterest*. Retrieved from <https://www.pinterest.com/pin/232076187020778284/?lp=true>

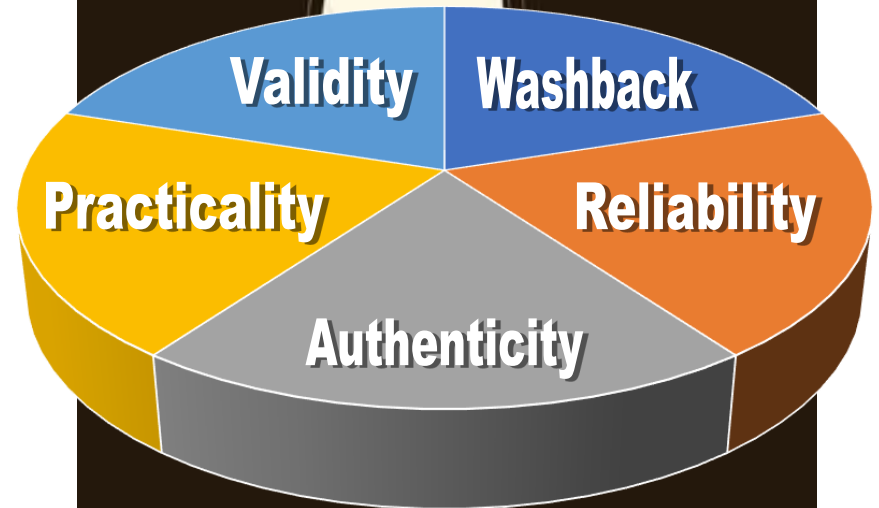
Reece, J. B., Urry, L. A., Cain, M. L., Wasserman, S. A., Minorsky, P. V., Jackson, R., & Campbell, N. A. (2014). *Chemical context of life. Campbell biology* (10th ed.). Boston: Pearson.

WIDA can do descriptors. (n.d.). *Board of Regents of the University of Wisconsin System*. Retrieved from <https://wida.wisc.edu/teach/can-do/descriptors>

Final DUE 04-30-19 Part 2

Assessment in ESL
CTES 7486
Dr. Jamie Harrison

Assessment Principles



PIE RECIPE BOOK

Created By Tabby Self

Practical PIE Pumpkin RECIPE



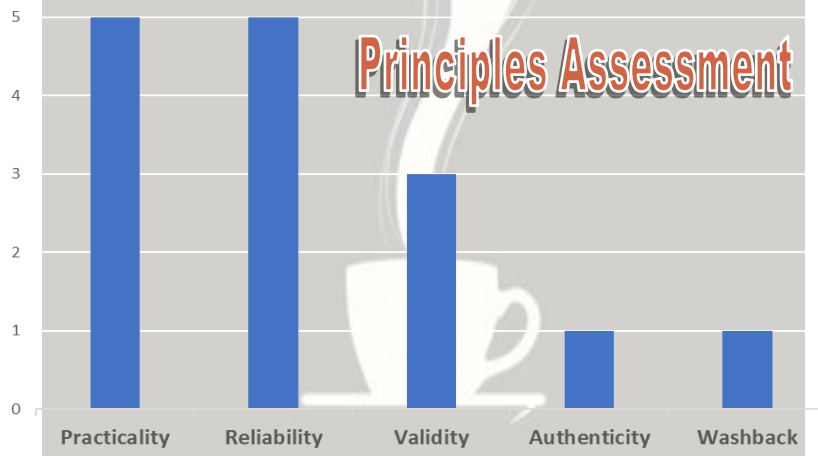
The principle of practicality is maximized when pragmatic logistics of assessment administration and design are prioritized. For example, considerations such as time constraints, budgets, resource and material availability, and scoring efficiency become the primary focus (Brown & Abeywickrama, 2010).

Sample Task

Circle the correct response from the choices below.

- Which of the following ships were not part of Columbus' maiden voyage to the New World?
 - Pinta
 - Niña
 - Mayflower
 - Santa Maria

Principles Assessment



Task Analysis

- Practicality:** Task design and administration is easy and cost effective. Working as partners provides relief from time-consuming scoring and feedback, but at the same time, delivers impactful formative information to students.
- Reliability:** The final products of this task will vary greatly but should result in fairly similar levels of formative impact. Reliability is less of a concern in such formative tasks.
- Validity:** Assuming a primary unit objective is to identify, understand, and produce key content terminology, the task reinforces this goal. Successful completion of the assignment provides a sampling of future performance and provides positive and non-critical insight into students' current knowledge.
- Authenticity:** Reproducing textbook terminology is not representative of real-world scenarios.
- Washback:** The primary purpose of this task is to provide students with insight into their current subject-matter knowledge in a non-judgmental manner and to encourage reflection on future summative assessment preparation.

ELL Modifications

Choose a paragraph from Chapter 2.



Copy paragraph leaving 5 blanks.

A _____ is a substance consisting of two or more different elements combined in a fixed ratio. Table salt, for example, is _____ (NaCl), a compound composed of the elements sodium (Na) and chlorine (Cl) in a _____ ratio. Pure sodium is a metal, and pure chlorine is a poisonous gas. When chemically combined, however, sodium and chlorine form an edible compound. _____ (H₂O), another compound consists of the elements hydrogen (H) and oxygen (O) in a _____ ratio. These are simple examples of organized matter having emergent properties. A compound has characteristics different from those of its elements (Figure 2.2). (Reece et al., 2014)

Exchange papers with a partner and fill in the blanks. Check the answers in your textbook and discuss with your partner.



WIDA Can Do Descriptors (n.d.) for 9-12th grade reading and writing state that level 1 ELLs are able to "supply missing words in short sentences," and level 2 ELLs are able to "follow multi-step instructions supported by visuals or data." Therefore, the instructions have been reformatted into simple commands with supporting illustrations and an example. Working in groups of two should provide any additional assistance needed for a beginner ELL to successfully complete this task.

Washback

Whoopie

PIE RECIPE



The principle of washback constitutes both positive and negative feedback. Assessments that encourage positive washback provide opportunities for formative feedback that can result in learning and teaching adjustments. Negative feedback typically emerges from score-focused and competitive assessments, often lacking formative information (Brown & Abeywickrama,

Sample Task

Choose a key paragraph from Chapter 2 of our biology textbook. Rewrite the paragraph substituting blanks for five important words or phrases. Exchange papers with a partner and complete the paragraph as best as possible without referring to your textbook. Afterwards, open your textbook to Chapter 2 to check answers and complete any missing information. Discuss your findings with your partner.

Principles Assessment



Task Analysis

- *Practicality*: Multiple-choice tasks are the quintessential practical assessment tool. The tasks are easy to design, administration is simple and low-cost, and scoring is straightforward and efficient.
- *Reliability*: Completing and scoring multiple-choice tasks are typically unambiguous, objective, and consistent, resulting in high test, rater, student, and administrative reliabilities.
- *Validity*: Assuming the objective of the task is determine whether or not the student can identify Columbus' three ships, content validity is present. However, the probability of guessing the correct answer on multiple-choice questions is reasonably high. False positives would incorrectly predict future performance.
- *Authenticity*: The task is a non-contextualized, isolated question. Test-takers are not likely to encounter real-world situations that pose questions or require responses in this format.
- *Washback*: Multiple choice tasks can result in negative washback because there is no "gray area" for students to demonstrate their knowledge—the response is 100% right or wrong. Multiple-choice outcomes provide little constructive feedback and are score-driven rather than formative.

ELL Modifications

Circle the correct response.

1. Which of the following ships were ~~NOT~~ part of Columbus' maiden voyage to the New World?
 - a. Pinta
 - b. Niña
 - c. Mayflower
 - d. Santa Maria



WIDA Can Do Descriptors (n.d.) for 6-8th grade reading state that level 1 ELLs are able to identify "responses to Wh- questions in charts or illustrated text." The added picture provides text illustration to support comprehension of complex language. The multiple-choice Wh- question and answer choices correspond to the Can Do Descriptors and require little modification. The words "not" and "circle" were emphasized to help ELLs zone in on critical instructional information.

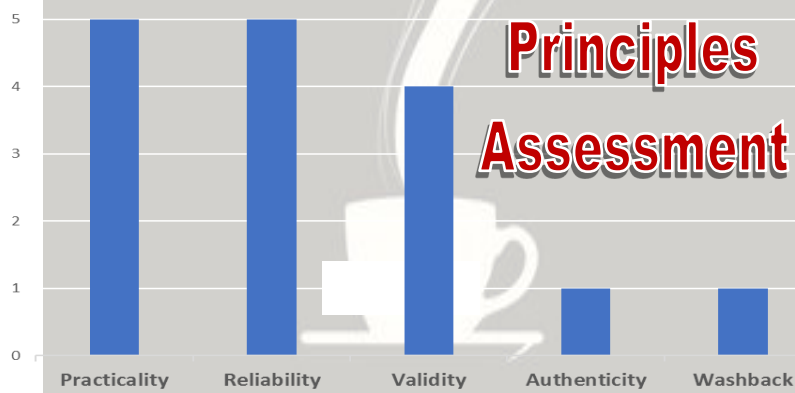
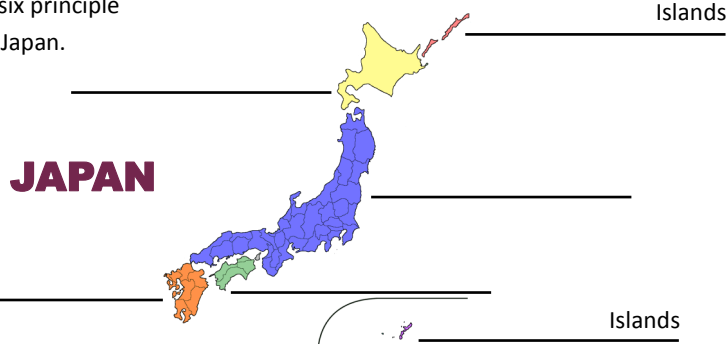
e
rb

A stylized illustration of a golden-brown pie with red filling, topped with two cinnamon sticks. The letters 'e' and 'rb' are visible on the left side of the image.

The principle of reliability is determined by the degree of consistent outcomes over multiple administrations. Reliability is applicable to the test design, rater, testing conditions, and test-taker. Assessment design can take into consideration each of these four areas, but some more realistically than others. For example, test design typically cannot predict variances in individual test-taker fatigue on a particular day, but test design can bolster test condition reliability by providing a climate controlled and quiet atmosphere (Brown & Abeywickrama, 2010).

Sample Task

Label the six principle islands of Japan.



Task Analysis

- *Practicality*: While the task is easy and cost-effective to design and administer, developing a scoring rubric and assessing the final product may be time consuming and difficult.
- *Reliability*: The subjective and open design of the task is likely to lead to a wide variety of final products and scoring inconsistencies across multiple administrations. A well-designed rubric that lays out definable grading criteria as well as inter-rater training and intra-rater awareness would help considerably with reliability concerns.
- *Validity*: If a lesson objective is to understand and apply knowledge of macronutrients to everyday situations, this task achieves content validity. The practical and meaningful use of acquired information is likely to result in positive consequential validity and enhanced retention that could positively predict future performance.
- *Authenticity*: The task's authenticity is maximized through the application of content knowledge under meaningful and natural conditions.
- *Washback*: The task provides positive feedback under realistic conditions. Students are likely to see content relevancy and thus find motivation to actively engage in classroom participation. The open format also gives students the freedom to personalize the project, promoting positive impact.

ELL Modifications

Fill out a
FOOD DIARY
for 7 days.

Answer below after 7 days.

Positive Comment:

Recommendation:

Food Diary		DAY 1		Daily Calorie Goal <u>2100</u>				
Meal	Food / Drink	# of Servings	Serving Size	Protein (g)	Carbs (g)	Sugar (g)	Fat (g)	Calories
Tue, Aug 15, 2017	L	1	1	10.3	42.2	12	14.2	327
165	DAILY TOTALS (Calories Remaining: 1823)			10.3	42.2	12	14.2	327

Carb-Fat-Protein Ratio: ____g : ____g : ____g

Carb %: % (total daily carbs / total daily calories)

Fat %: _____ % (total daily fat / total daily calories)

Protein %: _____ % (total daily protein / total daily calories)

Sugar-Carb Ratio: ____g : ____g

(Food Diary Template, n.d.)

WIDA Can Do Descriptors (n.d.) for 9-12th grade reading and writing state that level 2 ELLs are able to “classify or organize information presented in visuals or graphs” and “make content-related lists of words, phrases, or expressions.” Therefore, the task instructions and design were reformatted into a graphic organizer with an example to help scaffold instructions and writing production.

Authentic PIE RECIPE Apple

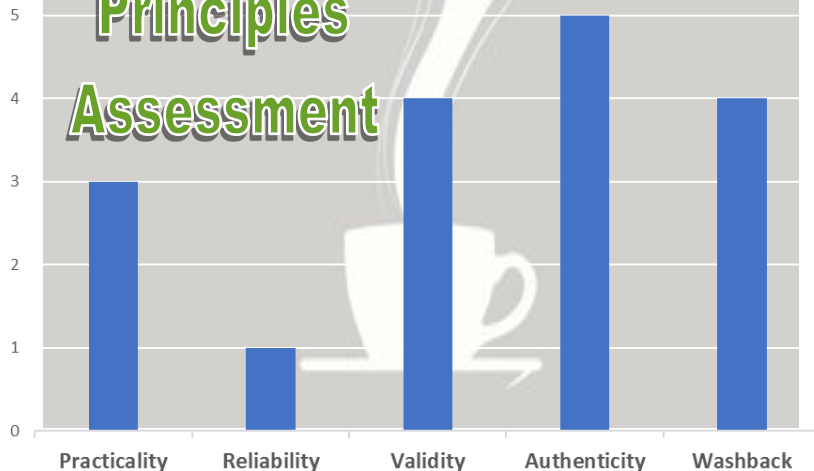


The principle of authenticity prioritizes the application of knowledge to real-world situations and the meaningfulness of tasks to everyday lives. Authenticity can be achieved through contextualization and reproduction of natural conditions (Brown & Abeywickrama, 2010). When information is interesting and relevant, students become more motivated to extrapolate knowledge to their personal surroundings thus enhancing knowledge retention.

Sample Task

Our food is composed of three primary macronutrients: Carbohydrates, fats, and protein. Design a daily tracker and document the food you consume for one week, noting the carbs, fats, protein, sugar, and calories. For each day, total each category and calculate your carb-fat-protein ratio, the percentage of carbs, fats, and protein to total calories, and your sugar-carb ratio. What positive findings did you notice? What changes to your diet would you suggest?

Principles Assessment



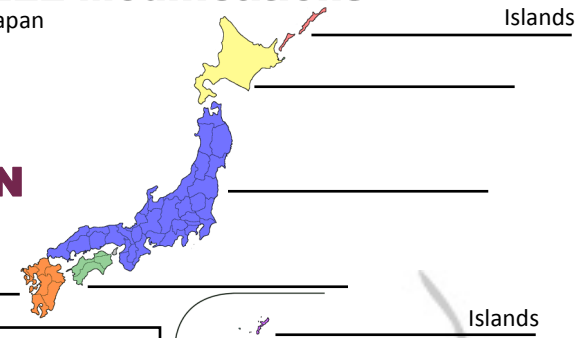
Task Analysis

- *Practicality*: Labeling tasks are very practical. They are easy to design, score, and administer on a low budget and within tight time constraints.
- *Reliability*: Labeling tasks exhibit high reliability qualities because instructions are clear, completion is straightforward, scoring is objective, and administrative conditions can be easily controlled.
- *Validity*: Assuming the objective of the task is to determine whether or not the student can identify the major island groups of Japan, validity is relatively high. Little probability of guessing correct responses is present, making this task more indicative of content knowledge and future performance.
- *Authenticity*: The task is a non-contextualized, isolated question. Test-takers are not likely to encounter real-world situations that pose questions or require responses in this format.
- *Washback*: Labeling tasks can result in negative washback because there is little “gray area” for students to demonstrate their knowledge. Labeling task outcomes provide little constructive feedback and are score-driven rather than formative.

ELL Modifications

LABEL the 6 islands of Japan using the **WORD BANK**.

JAPAN



WORD BANK

Honshu	Kyushu	Kuril
Okinawa	Hokkaido	Shikoku

WIDA Can Do Descriptors (n.d.) for 6-8th grade reading and writing state that level 1 ELLs are able to match “content-related objects, pictures, or media to words and phrases” and reproduce “words and phrases related to topics.” A word bank was added to assist beginner ELLs in content vocabulary production and thus more accurately reflect content knowledge. Accurate guessing could, however, reduce validity. To help interpret instructions, important words were emphasized even though the task is rather self-explanatory.

Validity

PIE Divinity

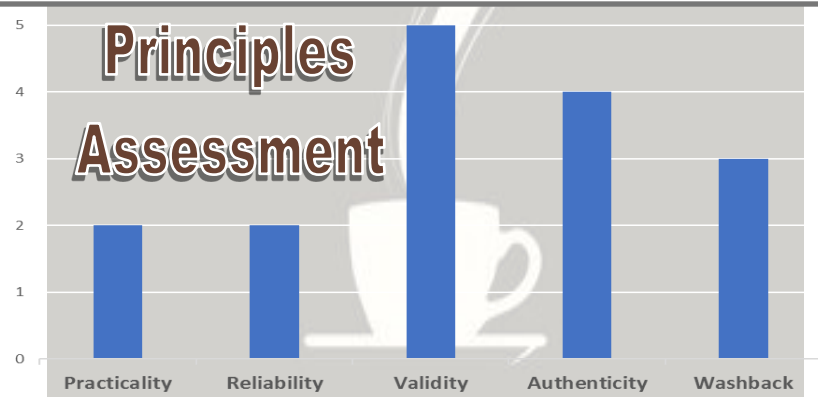
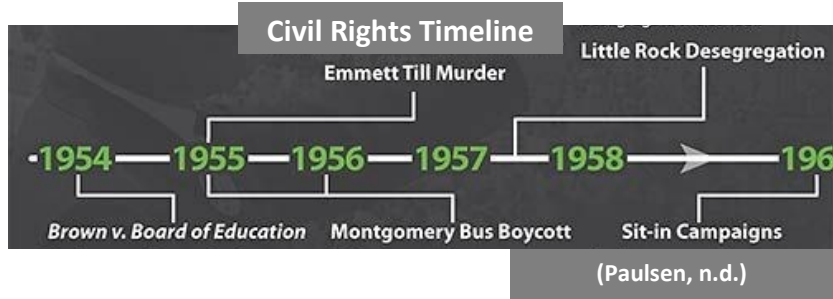
RECIPE



The principle of validity is concerned with how well an assessment aligns with unit objectives, whether the assessment results are predictive of future performance, and how the assessment impacts test-takers. Valid assessments tend to be performance-based and involve more authentic tasks resulting in meaningful and formative outcomes (Brown & Abeywickrama, 2010).

Sample Task

Write a short description, 3-5 well-formed and detailed sentences, for each of the six events listed on the Civil Rights Timeline below. Include significant people and their roles, primary locations, and descriptions of major events and outcomes.



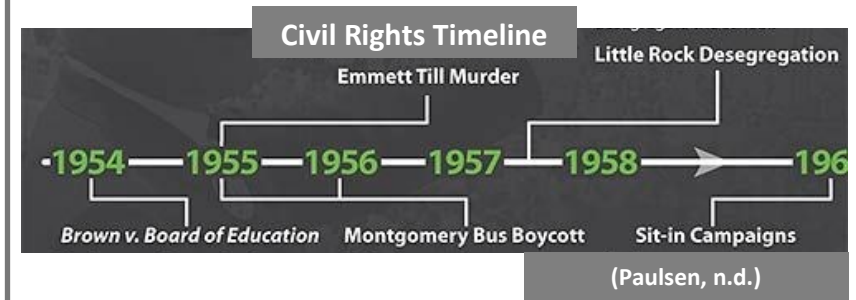
Task Analysis

- **Practicality:** While the task is simple to create, the scoring is quite complex. Open-ended, subjective tasks require documented scoring parameters, such as rubrics, inter-rater training, and time-consuming grading that is subject to rater variabilities.
- **Reliability:** Open-ended, subjective tasks are likely to result in inconsistent outcomes across multiple administrations. Reliability can be somewhat controlled with explicit instructions and scoring parameters.
- **Validity:** Assuming a primary unit objective is to understand the content of the major Civil Rights events as outlined on the timeline, successful completion of the task would likely correlate highly with future performance in this subject area. The formative nature of the assignment and opportunity to display knowledge under low-stress conditions is likely to result in positive tester impact.
- **Authenticity:** The task is contextualized into a timeline, meaningful in content, and interesting.
- **Washback:** Students have the opportunity to demonstrate knowledge and receive formative feedback from scorer comments..

ELL Modifications

Fill in the following information for each of the **6 events** on the timeline below:

Who: _____ (list people)
 Where: _____ (list locations)
 What Happened: _____ (1 sentence)
 Conclusion: _____ (1 sentence)



WIDA Can Do Descriptors (n.d.) for 6-8th grade reading and writing state that level 2 ELLs are able to identify “how content-related phenomena relate to one another in illustrated text or media” and connect “short sentences.” The original task was broken down into more manageable, scaffolded units. ELLs could utilize non-classroom time and translators to complete the task successfully, and level 1 ELLs may partner with more proficient English-speaking students.