



- ✓ What evidence, if any, were you given about the validity of the test?
- ✓ Did you sense any biases in the test?
- ✓ What aspects of your language proficiency were not addressed but might have been?
- ✓ Were the results of the test useful and informative? For what purposes?
- ✓ Did the test influence, either positively or negatively, how you will study or learn the language?
- ✓ Were the results of the test reliable? Do you think you would have received the same results on a different day, with different content or assessors, or in a different location? Why or why not?
- ✓ What could have been done to make them more reliable?
- ✓ Were the test practical? Was the amount of time and effort you spent doing the test reasonable, appropriate to your purposes, and worthwhile?
- ✓ What could have been done to make the test more practical for your purposes?

### Step 3 ( † † )

Consider the differences between high-stakes and low-stakes assessments. That is, stakes from a test are high if they have major consequences on a person's life, such as entry to or completion of an educational program, hiring for employment, or immigration decisions. In contrast, stakes from many routine classroom assessments may be low if they are informal, do not greatly influence students' grades, or just involve observations or responses. What differences in expectations are there for validity, reliability, and practicality between high-stakes tests and low-stakes assessments? What evidence may be required in each type of assessment situation to demonstrate validity and reliability? Make a list of differences, factors, and issues.

### Step 4 ( † † † )

Share the ideas you have discussed with other pairs. Make a table of these ideas, and suggest how they could be applied in classroom assessments.

