1996

Validity Issues in Computer-Assisted Strategy Assessment for Language Learners

Carol Chapelle

Iowa State University, carolc@iastate.edu

Follow this and additional works at: http://lib.dr.iastate.edu/engl_pubs

Part of the Bilingual, Multilingual, and Multicultural Education Commons, Curriculum and Instruction Commons, Educational Assessment, Evaluation, and Research Commons, and the Educational Methods Commons

The complete bibliographic information for this item can be found at http://lib.dr.iastate.edu/engl_pubs/56. For information on how to cite this item, please visit http://lib.dr.iastate.edu/howtocite.html.

This Article is brought to you for free and open access by the English at Iowa State University Digital Repository. It has been accepted for inclusion in English Publications by an authorized administrator of Iowa State University Digital Repository. For more information, please contact digirep@iastate.edu.
Validity Issues in Computer-Assisted Strategy Assessment for Language Learners

Abstract
This article probes validity issues of computer-assisted strategy assessment (CASA) in second-language (L2) acquisition research. Using two CASA studies, it reviews the metacognitive and cognitive strategies. It discusses trait-oriented and interactionist approaches to the concept of strategy, and examines validity by focusing on "advance preparation" and "resourcing," examples of trait-oriented and interactionist strategies, respectively. It suggests additional forms of validity evidence and describes applications of CASA as well as problems associated with this form of measurement.

Disciplines
Bilingual, Multilingual, and Multicultural Education | Curriculum and Instruction | Educational Assessment, Evaluation, and Research | Educational Methods

Comments
This article is from Applied Language Learning 7 (1996): 47. Posted with permission.
Validity Issues in Computer-Assisted Strategy Assessment for Language Learners

Carol A. Chapelle
Iowa State University

This article probes validity issues of computer-assisted strategy assessment (CASA) in second-language (L2) acquisition research. Using two CASA studies, it reviews the metacognitive and cognitive strategies. It discusses trait-oriented and interactionist approaches to the concept of strategy, and examines validity by focusing on "advance preparation" and "resourcing," examples of trait-oriented and interactionist strategies, respectively. It suggests additional forms of validity evidence and describes applications of CASA as well as problems associated with this form of measurement.

Researchers of L2 acquisition and computer-assisted language learning (CALL) have exploited computer technology as means of gathering performance data of learners working on language activities. These data are then used to make inferences about learners’ linguistic competence (e.g., Bland, Noblit, Armington, & Gay, 1990) and strategies for L2 acquisition. As computer technology becomes more widespread throughout educational settings, these research practices are likely to expand in their popularity. In anticipation that more researchers will be interested in CASA for L2 learners in the future, this article draws upon past experience to summarize pertinent issues.

CASA Research

Throughout the 1980’s and early 1990’s, a number of researchers have used computer-assisted methods for strategy research by collecting data generated by learners while using interactive software. In such studies, the recorded data were generated through the learner input to the computer program and the program responses to the learner. This work has not been reexamined and interpreted from a measurement perspective to identify common issues across individual studies and to define "CASA" as a method.

Although there are many possible roles for the computer in the learning environment and therefore in strategy research (e.g., Chun, 1994), the basic learner-computer interaction that occurs during individualized work has formed the basis for most L2 strategy research to date. I will focus on assessment of strategies familiar from other L2 research rather than the fundamental cognitive processes that have been the object of inquiry in psychology laboratories for many years.¹

When a computer program is used to control or facilitate a language task, it can unobtrusively document learners’ behavior such as the time they spend at various points in a problem sequence, the order in which they complete steps, and the editing they do to produce a linguistic product. Such a program can collect as sequenced data everything the learner sees on the screen, all keypresses and mouse-clicks that the learner...
makes, and the time that each action takes place. For the most part, software in such research has been
designed or modified to capture and store the desired information. The question for the researcher is which of
the obtainable data can be used as indicators of learner capacities. The CASA studies listed in Table 1 provide
some tentative answers to this question.

One study of English as a second language (ESL) learners (Jamieson & Chapelle, 1987) investigated "advance
preparation," a metacognitive strategy defined as "planning for and rehearsing linguistic components necessary to
carry out an upcoming language task" (O’Malley et al., 1985, p. 33). Using ESL dictation tasks over the course
of six weeks, learners listened to words (on the "spelling" task) and individual sentences (on the "dictation" task)
and then typed what they had heard. In this setting, it was assumed that the time the learner spent after hearing
the input and before responding was spent planning performance; therefore, advance preparation was inferred
from the amount of time that elapsed between the end of the input signal and the time that the learner pressed
the first key to begin to answer. The actual indicator of the degree to which each learner used advance
preparation was obtained by having the computer store the time it took to respond to each item, and calculating
the mean "time-to-begin" by dividing each learner’s total by the total number of items she or he had completed.
It is important to note that the learners in this study were not instructed to complete the exercise as quickly as
possible; they were working routinely over the course of several weeks.

Another metacognitive strategy, monitoring output, was assessed in the same instructional setting. On the
dictation tasks, learners were able to edit the response they had typed by deleting, inserting, and changing
characters or words before the response was evaluated by the computer. This behavior was documented in the
computer records; it was considered an indicator of monitoring output, in the sense described by Krashen
(1982), Bialystok (1981), Wenden (1985), and O’Malley et al. (1985). The number of times a learner edited
was divided by the total number of completed items to obtain the average number of times a learner edited each
item.

A third metacognitive strategy investigated with the same software was "monitoring input," defined by Bialystok
(1981) as, reflecting on the formal aspects of a message as it was comprehended. The dictation tasks allowed
learners to listen to the input as many times as they wanted before typing it. The demands of the task required
learners to focus on formal aspects of the input. When they had not comprehended a sentence or word the first
time it was presented, they requested to hear it again one or more times. Subsequently, those requests were
used as evidence for monitoring input. In other words, "monitoring input" was inferred on the basis of observed
requests.

<table>
<thead>
<tr>
<th>Assessment context</th>
<th>Learner</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Task goal</strong></td>
<td><strong>Behavior</strong></td>
</tr>
</tbody>
</table>

**TABLE 1**
*Studies using computer-assisted strategy assessment*
<table>
<thead>
<tr>
<th>Task</th>
<th>Action</th>
<th>Instruction</th>
<th>Strategy</th>
</tr>
</thead>
<tbody>
<tr>
<td>construct &amp; edit sentences</td>
<td>completing extra sentences and</td>
<td>instruction</td>
<td>exploration</td>
</tr>
<tr>
<td>construct &amp; edit sentences</td>
<td>experimenting any responses</td>
<td>instruction</td>
<td>resourcing</td>
</tr>
<tr>
<td>read &amp; answer questions or summarize</td>
<td>requesting dictionary</td>
<td>experiment definitions</td>
<td>resourcing</td>
</tr>
<tr>
<td>complete grammar exercise</td>
<td>requesting rule/example help</td>
<td>instruction</td>
<td>inferencing/hypothesis-testing</td>
</tr>
<tr>
<td>complete grammar exercise</td>
<td>requesting/not requesting help</td>
<td>instruction</td>
<td>controlled/automatic processing</td>
</tr>
<tr>
<td>dictation: listen &amp; write word/sentence</td>
<td>latency before responding</td>
<td>instruction</td>
<td>advance preparation</td>
</tr>
<tr>
<td>dictation: listen &amp; write word/sentence</td>
<td>editing response</td>
<td>instruction</td>
<td>monitoring output</td>
</tr>
<tr>
<td>dictation: listen &amp; write word/sentence</td>
<td>requesting/repeated input</td>
<td>instruction</td>
<td>monitoring input</td>
</tr>
</tbody>
</table>

1 Hsu, Chapelle, & Thompson (1993)
2 Chapelle & Mizuno (1989)
3 Hulstijn (1993)
4 Doughty (1987)
5 Jamieson & Chapelle (1987)

* This refers to the learner's perception of what he/she is doing while participating in the activity. The learner's perception is believed to affect performance, and therefore should be considered in the interpretation of the performance. For example, "instruction" means the learner perceived the task as taking place for instructional purposes.

Other interaction observation programs have been used to measure "resourcing" defined by O’Malley, et al. (1985) as the cognitive strategy of using reference materials to obtain information about the L2. Consistent with this definition, researchers have considered each request for on-line help as evidence for the use of resourcing. In one study (Hulstijn, 1993), learners had access to an on-line dictionary that they could consult while reading a passage followed by questions. The resulting data documented the words that they had looked up in the dictionary and the time that each was looked up. Also while investigating resourcing, Chapelle and Mizuno (1989) collected data as learners worked on a task requiring them to construct and edit sentences in the exercises. As learners worked, help was given only at their request; they could ask for help with vocabulary, grammar, or the semantic facts pertaining to the sentences.

Doughty and Fought (1984) also documented learners' help requests on grammar items but interpreted them differently. Learners' requests for help while they were used as indicators of learners' "controlled access of explicitly learned knowledge." In addition, the researchers reasoned that "attempts to complete tasks without any help from the program [they reasoned] reflect automatic access to implicit knowledge in memory" (Doughty, 1987, p. 151). Other strategies evidenced by learners using particular types of help in software were
"hypothesis-testing" and "inferencing based on L1." Doughty and Fought operationalized the definitions of these strategies as the type of grammar help requested by learners. When learners consulted help consisting of examples, they were considered to be displaying evidence of "hypothesis-testing." When they chose grammar help consisting of formal rules or the correct answer, learners were considered to be "inferencing [about syntactic forms in the L2] based on L1."

Using a sentence-constructing and -editing task, Hsu, Chapelle and Thompson (1993) assessed another strategy—exploration, the use of software to experiment and test hypotheses about the target language (Higgins & Johns, 1984). Exploration was operationalized in two ways: the number of sentences learners constructed after having completed the number required by their assignment and the number of times they edited an answer after receiving a message that it was correct—which the software allowed but did not require them to do. The operational definition of exploration in this study was derived from the theoretical definition that included the notion that learners would be motivated and interested in experimenting with the software.

In summary, a number of strategies have been investigated using computer-assisted methods. To understand CASA as a method, however, it is informative to interpret the types of strategies from a measurement perspective and, consequently, to explore the nature of the strategy construct.

**Definition of "Strategy" in Computer-Assisted Strategy Research**

To examine this method of assessment, I consider strategies to be theoretical constructs that are themselves not observable, but are hypothesized to be responsible for observed behavioral data. Since there is more than one way to define a theoretical construct, I examine carefully how a particular strategy such as "monitoring" is defined in a research study. By looking for similarities in strategy definition across computer-assisted research studies, I query the extent to which an approach to strategy definition is tied to CASA. I provide means for considering the "nature of the construct," by distinguishing two approaches to construct definition. I then illustrate computer-assisted strategy research that has taken each approach to construct definition, thereby demonstrating that the method of measurement does not preclude either one.

I will distinguish between two approaches defining a strategy as a theoretical construct (approaches explained in Chapelle, forthcoming). The first is a trait-oriented definition that conceptualizes a strategy as an attribute of an individual independent of the context in which it is observed. One thinks of a strategy as a trait when one talks about "monitoring" as something learners do all of the time regardless of whether they are listening to an academic lecture, writing an e-mail message to a colleague, or speaking to a close friend. A trait-oriented construct definition assumes that a researcher is able to generalize the inferences made about a construct on the basis of performance on an assessment (i.e., performance in one context) to inferences about the construct in other contexts. In other words, if an individual is a strong monitor user in a test of monitoring, the trait definition would assume that the individual would also be a strong monitor user in the other contexts, such as instructional settings.

A second and contrasting way of defining a strategy as a theoretical construct is an interactionalist definition. This definition presents a strategy as a context-dependent attribute of an individual. From an interactionalist approach, one could not define "monitoring" in a global sense. Instead, one would refer to "monitoring while listening to academic lectures," for example. The definition of the strategy would include the context in which the strategy is used. To interpret results of a test of "monitoring while listening to academic lectures," the
researcher would generalize results only to monitoring in this context. In short, an important distinction between the two approaches rests on how far the strategy definition assumes the researchers can generalize the results of strategy assessment.

Both approaches to strategy definition have been used in computer-assisted strategy research. The assessment of "advance preparation" illustrates a trait-oriented definition in computer-assisted strategy research (Jamieson & Chapelle, 1987). The strategy is defined in a general way; even though the definition mentions the word "task," it does not refer to any particular task, implying that the strategy is conceived as one that could apply equally to a linguistic task in any context. Table 2 summarizes the key measurement facets of this strategy.

**TABLE 2**
*Examples of trait and interactionalist approaches to definitions of L2 strategies*

<table>
<thead>
<tr>
<th></th>
<th><strong>Trait-oriented definition</strong></th>
<th><strong>Interactionalist definition</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Example</td>
<td>advance preparation (Jamieson &amp; Chapelle, 1987)</td>
<td>resourcing (Chapelle &amp; Mizuno, 1989)</td>
</tr>
<tr>
<td>Definition</td>
<td>&quot;planning for and rehearsing linguistic components necessary to carry out an upcoming linguistic task&quot; (O’Malley, Chamot, Stewner-Manzanares, Kupper, and Russo, 1985)</td>
<td>&quot;use of target language reference materials” (pp. 28-29) in the context of learner-controlled CALL materials (p. 26)</td>
</tr>
<tr>
<td>Measurement</td>
<td>the amount of time (to .5 second accuracy) between the time that a prompt was given (in a CALL activity) and the time that the student began to respond (averaged over the number of items that student responded to over the course of the semester)</td>
<td>the frequency of the number of requests for help a student made divided by the number of sentences the student produced in a sentence constructing and editing CALL activity (help=dictionary, semantic/pragmatic facts, and grammar)</td>
</tr>
<tr>
<td>Inference</td>
<td>performance was assumed to indicate the degree to which the learner was an &quot;advance preparer&quot;</td>
<td>performance was assumed to indicate the degree to which learners used resourcing within the learner-controlled CALL activity</td>
</tr>
<tr>
<td>Use</td>
<td>to investigate the relationship between advance preparation and cognitive style as well as the relationship between advance preparation and subsequent language proficiency</td>
<td>to investigate the extent to which learners use resourcing in a set of learner- controlled CALL activities for practicing grammar and editing (as a means of evaluating the pedagogical potential of optional help)</td>
</tr>
</tbody>
</table>

Advance preparation was measured through response latency in an instructional setting in which learners were working at their own pace. The inferences made on the basis of summed response latencies were the degree to which the language learners were advance preparers in general, rather than the degree to which they used...
advance preparation while working on this type of software. Accordingly, the scores for advance preparation were used in this research context to investigate the relationship between this strategy and other variables, which were also defined in a context-independent manner.

The assessment of "resourcing" provides an example of an interactional approach to strategy definition (Chapelle & Mizuno, 1989). It is defined in this research as a learner’s use of target language reference materials in learner-controlled CALL materials. The definition is interactionalist because it includes the "learner-controlled CALL materials" as the context to which we wish to generalize. Measurement of the strategy was calculated by tabulating the number of times the learners requested help per unit of activity (defined by construction of one sentence). The inferences made were intended to be limited to contexts of learner-controlled CALL, and the scores were used to evaluate the value of offering learners optional help in learner-controlled CALL.

**Justifying Inferences about Strategies from Observed Performance Data**

In the research cited above, as in any research, one is ultimately concerned about the validity of the uses that are made of the strategy assessment. Validity of test use rests on justifying the inferences made from observed behaviors. Justifications of inferences about strategies constitute evidence for the construct validity of those inferences. Table 3 summarizes and defines types of construct-validity evidence suggested by Messick (1989).

To examine these types of construct-validity evidence, I return to the examples of advance preparation and resourcing. Then I will suggest for each example additional evidence that could be used to make a stronger case for construct validity. Finally, I will underscore two fundamental points about construct validity: (1) the nature of the evidence depends on the way "strategy" is defined in the research and (2) construct-validity evidence refers to the justifications provided for interpretations and therefore it can be evaluated as a strong or weak relative to particular inferences rather than as an all or nothing quality of an assessment.

In the first study, Jamieson and Chapelle (1987) provided three types of justifications that might be used to argue for inferring "advance preparation" from response-latency data. First, content-validity evidence consisted of the authors’ judgment that this behavior in the instructional context logically fit with the definition of the construct:

\[\text{Advance preparation was inferred from the amount of time it took for the student to press the first key of his or her answer. O’Malley et al. (1985) defined advance preparation as a metacognitive strategy that means "planning for and rehearsing linguistic components necessary for an upcoming language task" (p. 33). The student behavior of consistently waiting before answering may indicate the degree to which he or she was engaging in preparation to answer (p. 531).}\]

**TABLE 3**

*Potential methods for justifying construct-validity of inferences from tests/assessments (based on Messick, 1989)*

<p>| Content analysis | Experts’ judgments of what they believe a test measures |</p>
<table>
<thead>
<tr>
<th><strong>Empirical item investigation</strong></th>
<th>Investigation of factors affecting item difficulty and discrimination to provide statistical evidence relevant to researchers’ understanding of what the test measures</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Internal consistency</strong></td>
<td>Estimation of the consistency of learners’ performance</td>
</tr>
<tr>
<td><strong>Empirical task analysis</strong></td>
<td>Documentation of the metacognitive strategies that learners use as they complete test tasks through qualitative methods such as think-alouds</td>
</tr>
<tr>
<td><strong>Correlations with other tests and behaviors</strong></td>
<td>Identification of theoretically predicted levels of covariance and behaviors among tests and behavior in appropriate contexts</td>
</tr>
<tr>
<td><strong>Experimental studies</strong></td>
<td>Investigation of changes in performance which accompany systematic changes in test design</td>
</tr>
</tbody>
</table>

* See Chapelle (1994) for an example of how these approaches to construct validity are applied to the evaluation of inferences and uses made from a test.

Second, the authors found a sufficient degree of consistency in strategy use over a six-week period to provide evidence that a construct was assessed through the observed data. "Mean working-style scores from six randomly selected weeks were treated like 6-item scales on which ... reliability estimates were made" (a=.72 and .82 on the two activities, p. 535). Third, a theoretically predicted correlation with a style variable "reflectivity-impulsivity" was found. Advance preparation was significantly, positively related to reflectivity (r=.50; p<.001); one would expect that this strategy "would logically be associated with the slow, careful aspect of the reflective learner" (p. 538).

These three arguments provide some justification for the advance preparation inferences made from learner performance data. However, the argument would be stronger if additional sources of evidence had been provided. First, ideally, evidence consisting of learner verbal reports might indicate that they were thinking about and planning what they would type between the time they heard the input and the time they began to respond. Second, the authors might also have provided correlations not only with a measure of a related construct but also with another measure of the same construct, "advance preparation." Because of the trait-oriented definition of advance preparation, the other measure should assess advance preparation through a different method of measurement to ensure that performance is the same regardless of the context in which it is measured. Third, some form of experimental data could also contribute to the validity argument. For example, an experiment might compare response latency data of subjects who had been trained to stop and plan with those who were told not to think before responding. If performance reflected the expected patterns for the two groups, results could be used as justification for making inferences about advance preparation from performance.

In the second example, the justification for Chapelle and Mizuno’s (1989) use of performance data to assess resourcing rested solely on content evidence for validity. The authors used the following justification: "The computer provides help only upon request so learners must ask for the help they need when they need it. Learners’ requests for help are [therefore] evidence of their use of resourcing..." (p. 28). This provides only one argument for their strategy interpretations.
Other justifications that would have made their argument stronger would include the following: First, they might have consulted learners’ verbal reports indicating that they had chosen help in order to obtain assistance with the sentence-construction task, rather than for other purposes (e.g., to see what the help looked like). Second, they could have demonstrated consistency in the use of help over the several weeks the activity was used. Third, they might have supplied correlations between use of help on their learner-controlled CALL activity and another one. To act as correlational validity evidence for the interactionalist definition of the strategy, the covariate must be similar to the original assessment in terms of assessing resourcing in learner-controlled CALL as well. Fourth, they could have used an experimental study comparing subjects who had been trained to use help with those who were not told to use it.

These examples indicate that the nature of the construct-validity evidence is related to the construct definition, and where correlational evidence is concerned the type of construct definition (trait-oriented vs. interactionalist) impacts the choice of covariate in validation research. Further, they also indicate that the legitimacy of inferences made from the data is a matter of degree rather than an all or nothing proposition. Validity justification consists of an argument relevant to understanding the meaning of observed data for making inferences about strategies—an argument that is essential for justifying the use of these assessments in SLA research.

**Applications of Computer-Assisted Strategy Research**

Although very few research projects have used CASA, the method has shown promise in several ways. First, CASA has been useful in extending the researcher’s ability to document behaviors in language-learning contexts particularly when large amounts of precise data must be tabulated to make valid inferences about strategies. For years, classroom researchers attempting to study learner strategies have been frustrated by the amount of relevant performance data that they were able to obtain through observation of behaviors (e.g., Naiman, Frohlich, Stern, & Todesco, 1978). Consequently, computer-assisted assessment offers a useful addition to strategy assessment methods.

Second, CASA allows for gathering strategy data during actual instructional exchanges. Given our increasing appreciation of the effects of learners’ perception of the task on their performance (e.g., Bruner, 1990), researchers increasingly value data that are obtained in genuine instructional contexts. Despite the usefulness of data obtained through obtrusive methods such as think-aloud protocols (Ericsson & Simon, 1984), and retrospective self-reports (Cohen & Hosenfeld, 1981), there is a need to complement them using observation programs that document behavior unobtrusively in classroom settings. Thus, researchers begin to integrate SLA and pedagogical research and strengthen the possibility of identifying strategies that might provide useful feedback to learners (Scott & New, 1994).

**Problems with Computer-Assisted Strategy Research**

I would characterize the current problems of computer-assisted strategy research as consisting of two types: analytic measurement problems and practical problems. Measurement problems refer to the need to investigate validity from the perspectives of both construct inferences and consequences of CASA use.
Methods for justifying construct validity are outlined above. As I illustrated, the validity of the inferences and uses of assessment needs to be justified. Because of the precision, accuracy, and directness of data collection in computer-assisted strategy research, however, researchers might attempt to argue away the need for construct validity justification as Chapelle and Mizuno (1989) did: "The behavior exhibited for [resourcing] provides unequivocal evidence for students’ use of [this strategy]; in a sense, the behavior is the strategy” (p. 34). From the perspective of the interactionalist construct definition for the resourcing strategy, the behavior is not equivalent to the strategy itself. A second construct validity problem arises when justifications are inappropriate for the type of inferences and uses made from the assessments. For example, if Chapelle and Mizuno (1989) had, on the basis of their content-validity argument, used their data to make inferences about the extent to which learners were resourcers across contexts and had used that data to decide who needed training in resourcing, these inferences and uses would have been inappropriate.

Consequential validity refers to justifications for the usefulness of an assessment for its intended purposes as well as for its unintended outcomes beyond the immediate assessment event and context. Potential consequential validity problems arise when learners are disturbed, rather than facilitated through the assessment process. For example, data gathered and used in a way that violates learners’ rights to privacy would argue against consequential validity, as would computer-assisted assessment of learners who feel uncomfortable using the computer. The potential detrimental effects of computer-assisted language learning in general have not been explored rigorously; however, studies in critical pedagogy (e.g., Bowers, 1988) provide some useful directions that may also pertain to the consequential validity of strategy assessment.  

Practical problems in CASA are also worth noting. First, in constructing a computer-assisted assessment, researchers may find it difficult to identify software that simultaneously provides relevant language learning activities for instruction and strategy assessment. Unfortunately, to date little work has attempted to combine efforts in instructional design with those of assessment of either strategies or language. A second practical hurdle for CASA is the challenge of modifying software to get it to gather the appropriate data. Some of the commercial software contains data collection capabilities, but there is no guarantee that a given piece of software will collect the data of interest to the researcher. Many of the CASA projects to date have been conducted by researchers who developed their own software. A third practical problem becomes the management of the large quantities of process data that can be generated by recording the details of learners’ interaction. Because disk space is limited, there is always a need to make rational decisions about how to summarize and store the data throughout the assessment process (Goodfellow & Laurillard, 1994). Past research has shown that while these practical problems present significant challenges, they are not insurmountable.

Conclusion

Research on interaction-observation programs has been found useful for assessing some SLA strategies. Methodologically, it is particularly interesting that the researcher can construct the type of learning environments which learners would use routinely for instructional activities but which simultaneously serve as a laboratory for data collection (Doughty, 1992). Moreover, the capability to investigate longitudinally learners’ routine "working styles" (Jamieson & Chapelle, 1987) offers an ideal setting for investigating important questions about learners’ strategies. For example, how do learners’ strategies change as task demands are manipulated or as they accrue experience with a task? How consistent (reliable) is strategy measurement on the same task across different
occasions? Can the accuracy of self-report data be substantiated by observation of computer-documented protocols? The investigation of these and other strategy questions, however, relies on the validity of the measurement used to assess strategies and therefore rests in part on validity issues.

Notes

1 For years, psychologists (e.g., Sternberg, 1977; Snow, 1981) have used response-latency measures for assessing psychological processes in laboratory experiments. In second language acquisition similar research questions have been raised particularly by those approaching SLA from an information processing perspective (McLaughlin, 1987). Automaticity in language processing is hypothesized to be indicative of language knowledge efficiently stored for expedient retrieval; as an aspect of the target language is better learned, restructuring of knowledge occurs making access more automatic. In experimental settings, the amount of time subjects take to respond to a task has been used as an indicator of how automatic subjects are in the use of the linguistic knowledge necessary for performance on the task, or whether knowledge is implicitly or explicitly stored. For example, Hagen (1994) used a computer program to present items to subjects and to time their responses to make such inferences. Crucial to the interpretation of time-to-respond in this case was the learners’ understanding that they were to respond as quickly as they could—a task demand the researcher had to make clear to the subjects.

2 This is not the only way of viewing a strategy. Another is to define it in terms of the actual observed behaviors (see Cohen, this volume).

3 Of course, the context of the academic lecture could be defined in greater detail to state a more specific construct definition.

4 What is needed to better express the interactionalist construct definition is a more complete and systematic way of defining "context." Proposals for approaching this problem have been suggested by researchers in language testing (Bachman, 1990), language instruction (Skehan, 1992), and SLA research (Duff, 1993).

5 Bowers (1988) points out the need to examine the range of possible activities learners engage in through computer-assisted instruction in order to shed light on what students are missing by spending time at the computer. She also guides us to examine the cultural ideologies—such as the value Western societies place on information and individualization—inherent in our educational uses of technology.

References


**Acknowledgment**
I am grateful for comments from Roberta Abraham and an anonymous reviewer on earlier versions of this article.

Author

CAROL A. CHAPELLE, Professor, TESL/Applied Linguistics, 203 Ross Hall, Iowa State University, Ames, IA 50011. Specializations: computer-assisted language learning, language testing, SLA research.