2009

Individual Learner Differences in CALL: The Field Independence/Dependence (FID) Construct

Carol Chapelle
Iowa State University, carolc@iastate.edu

Trude Fraiser
Simon Fraser University

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 Methods Commons, and the Other Computer Sciences Commons

The complete bibliographic information for this item can be found at http://lib.dr.iastate.edu/engl_pubs/46. 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.
Individual Learner Differences in CALL: The Field Independence/Dependence (FID) Construct

Abstract
In Spring semester 2006, we conducted a study with 50 learners of German to investigate the use of a new measure (Cárdenas-Claros, 2005) for research on the field independent/dependent (FID) cognitive style and CALL use. After the measure was administered, students worked on a CALL program for German that logs interaction. This paper reports the reliability and item analysis of the FID-CALL measure in addition to its relationship with learners' behaviors in the CALL program. With the goal of understanding how FID intersects with satisfaction and success in learning through CALL, we suggest areas for future development of the measure.

Keywords
Measure of Field Independence/Dependence, Individual Learner Differences, Student Use of CALL Programs, German

Disciplines
Bilingual, Multilingual, and Multicultural Education | Curriculum and Instruction | Educational Methods | Other Computer Sciences

Comments
This article is from CALICO Journal 26 (2009): 246, doi:10.1558/cj.v26i2.246-266. Posted with permission.
Individual Learner Differences in CALL: The Field Independence/Dependence (FID) Construct

Carol A. Chapelle
Iowa State University

Trude Heift
Simon Fraser University

Abstract
In Spring semester 2006, we conducted a study with 50 learners of German to investigate the use of a new measure (Cárdenas-Claros, 2005) for research on the field independent/dependent (FID) cognitive style and CALL use. After the measure was administered, students worked on a CALL program for German that logs interaction. This paper reports the reliability and item analysis of the FID-CALL measure in addition to its relationship with learners’ behaviors in the CALL program. With the goal of understanding how FID intersects with satisfaction and success in learning through CALL, we suggest areas for future development of the measure.

Keywords
Measure of Field Independence/Dependence, Individual Learner Differences, Student Use of CALL Programs, German

Introduction
Computer technology presents an opportunity for individualization of language learning that does not exist in the classroom. An important area for computer-assisted language learning (CALL) research is to attempt to better understand the ways in which individual differences affect learners’ use of CALL and the benefits different types of learners may obtain from the use of such materials. Researchers in second language acquisition (SLA) agree that individual differences (IDs) such as cognitive style, learning strategies, and affective variables influence learners’ processes and success (Skehan, 1998; Dörnyei & Skehan, 2003). One promise of research on IDs is individualized learning materials, and the most realistic hope for such individualization is CALL. Recent research has investigated working memory and motivation in CALL tasks (Payne & Whitney, 2002; Ushida, 2005), but the potential of CALL for making use of ID research has not been realized primarily because acceptable measures of the relevant IDs are needed. This paper describes research investigating a new scale for IDs, field independence/dependence (FID), designed for research on CALL.

Field Independence/Dependence
Numerous studies have found moderate correlations between field independence/dependence (FID) and success in language learning (Hansen & Stansfield, 1981; Brown, 2006). FID refers to an individual’s preference for internal versus external guidance (Witkin & Goodenough,
A student with a field independent (FI) style tends to create his or her own structure in problem solving and is therefore an individualist rather than one who engages in interpersonal deliberations. This FI learner tends to attain the best success in classroom language learning according to much of the past research. The field dependent (FD) learner is the opposite: he or she tends to accept the structure and assistance provided by others and approaches problem solving in a more social or collaborative way (Jonassen & Grabowski, 1993). Figure 1 illustrates the different modes of thinking that characterize the FI and FD learner. The curved arrow in the top cell symbolizes the self-reliant, reflective learner who works through problems on his or her own. The double headed arrows in the bottom cell symbolize the interpersonal style of the FD learner who works through problems in a consultative manner with other learners or informants. The FD learner prefers to work with others, and seeks their guidance. Thus, the FI student relies on his or her own internal frame of reference whereas the FD student draws upon others.

Figure 1
FI Versus FD Learner

The research on FID is interesting because of the relatively consistent findings of a connection between this cognitive style and SLA. Recent reviews of IDs in SLA, however, are critical of the research for three major reasons, all of which are relevant to the study of IDs and CALL.

First, previous research on FID has been conducted primarily within the framework of seeking a better understanding of the "good language learner." Although this quest is theoretically interesting, one might argue that such a goal alone fails to address the reality that
all—not only good—language learners need to learn. Research predicated on the idea that some learners are inherently better than others may be of limited use for the design and use of CALL materials.

Second, FID as a construct of ID was formulated years ago, when psychologists tended to think of constructs as traits that would influence behavior across a range of contexts. In other words, if you are field independent while you are doing math problems, you will be field independent while you are at your Spanish conversation lunch. Today, that conception of a construct is still used for some things, but in many cases a more contextually delimited construct proves to be more useful. In Dörnyei and Skehan’s (2003) review of the ID research in SLA, they conclude that “what is needed is more evidence of educationally linked applications of such concepts [as FID].” They suggest that “if such evidence is forthcoming, style concepts may become more central in SLA once again” (p. 607). Accordingly, for the study of IDs in CALL, FID needs to be defined not as a trait, but as an interactionalist construct which entails the interaction between the trait and the context of CALL use (Chapelle, 1998).

The third issue concerns the validity of interpretations and uses of the Group Embedded Figures Test (GEFT) that has been used to measure FID in SLA studies and elsewhere. The test tasks require learners to find and outline a simple geometric shape within a more complex figure. In view of the theoretical FID definition, one can question the extent to which the GEFT tasks fully sample from all aspects of the construct with only one type of task. Chapelle and Green (1992) noted that the GEFT only measures one of the components of the cognitive style definition proposed by Witkin and Goodenough (1981): “It is apparent that the EFT measures only the restructuring ability component” (Chapelle & Green, 1992, p. 51). Moreover, the aspect of the construct assessed resembles an ability more than a style. Brown (2007) also pointed out this problem and suggested that the GEFT fails to assess how individuals work because people are scored on the number of correct responses they get. As a result, the higher learners score on the GEFT, the higher the tendency such learners display toward the FI trait, but “a low score does not necessarily imply relatively high dependence” (p. 87). Despite these long-standing problems with this measure, it remains in use in research, presumably because of the lack of alternatives.

**FID and CALL**

Despite the less-than-satisfactory measurement of FID in previous research, there are three studies that have found relationships between FID and CALL. Abraham (1985) reported that FI learners of ESL learned an advanced grammatical rule best in CALL materials that provided a deductive presentation, whereas the FD learners did better in CALL materials that presented examples. Chapelle and Jamieson (1986) showed that the FD learners had more positive attitudes toward the CALL component of their ESL program than the FI learners, even when motivation was controlled for. Finally, in examining the correlates of learning strategy use, Jamieson and Chapelle (1987) found that strategy use should be assessed in conjunction with cognitive style. Their study which examined relationships between working styles and cognitive styles of 33 ESL students working with computerized spelling and dictation lessons, indicated that FI was significantly related to monitoring input. In addition, the FI students tended to work quickly rather than to stop and think.

These studies of learning efficiency, attitudes, and strategy use hold potential for matching learners with appropriate instruction. The rules versus examples formats for grammar teaching, for example, might realistically be offered to learners studying grammar online. CALL materials such as those used in Chapelle and Jamieson’s (1986) study could be offered...
to the FD learners, and a parallel set developed for the FI learners less fond of that kind of instructional design. Jamieson and Chapelle’s (1987) study suggests that it may be useful to consider the amount of learner control and help options for different task types and learners in a CALL program.

What would a good instructional design look like for the FD learners? They performed well with the deductive (rule presentation) design, but the CALL part of the curriculum in Chapelle and Jamieson’s (1986) study contained many lessons with such rule presentation. Perhaps the instructional design that the students like is not one that is good for their learning. Perhaps there were other features of the CALL materials to which the learners were responding. If we attempt to go beyond these studies to make recommendations for practice, it is immediately apparent that much more research is needed. In the interest of moving forward with such research, the measurement issues associated with assessment of FID in CALL need to be solved; that is, a measure needs to be developed to assess a construct of FID in CALL. Such a measure needs to be a genuine assessment of style, rather than ability, and needs to ask learners to respond to tasks that refer to aspects of CALL use (Chapelle, 1995).

Cárdenas-Claros (2005) attempted such a measure of FID for CALL use, but her initial trial of the measure was inconclusive perhaps because of a small sample size in her study. To reassess the measure, we administered it to 50 learners of German who used the E-Tutor, a parser-based CALL program for German that logs student-computer interactions.

**Research Questions**

This study follows up the investigation of the FID-CALL assessment developed by Cárdenas-Claros (2005) with CALL users who were studying German. Specifically, it addresses the following questions:

1. Is the FID-CALL scale developed by Cárdenas-Claros (2005) a reliable measure of individual differences when used to assess FID in CALL for first-year university students of German?
2. Are there aspects of CALL design covered on the scale about which students have unanimous opinions?
3. Were individual differences observed in the way that students interacted with the E-Tutor CALL program?
4. What is the relationship between responses on the FID scale and observed working style behaviors in the E-Tutor program?

**METHOD**

The study relied on self-report survey data and computer logs showing students’ behaviors while using E-Tutor during the 15-week session of their normal classroom instruction. Quantitative methods of item analysis, descriptive statistics, and correlations were used to examine the data. Records of program use were recorded over the 15-week semester in order to obtain an adequate sample of behaviors. These behaviors were summarized and examined in the light of learners’ responses to the FID-CALL scale.
Participants
The participants were 50 first-year students of German as a foreign language studying at Simon Fraser University in British Columbia. Sixty-five students taking beginning-level German initially agreed to participate in this study. However, while all students completed the course, only 50 chose to complete the FID-CALL scale, which was optional.

Materials
The E-Tutor (Heift, 2001; Heift & Nicholson, 2001) contains an online bilingual dictionary and grammar aids in the form of context-sensitive inflectional paradigms as well as explanations on the grammar topics covered in each chapter. In addition, students have access to the correct answers for all exercises and can also skip exercises at any point. The program gathers data as students use it. Student access of help aids provided the quantitative data of interaction for the study. In particular, we recorded the number of times students looked up a word in the dictionary, accessed the grammar aids (inflectional paradigms and static grammar notes), looked up an answer for an exercise, and/or skipped an exercise after the system signaled a mistake.

The measure of FID consisted of a 30-item questionnaire (Cárdenas-Claros, 2005) with six main sections, including questions about vocabulary, grammar, reading, writing, and listening, as well as a section on general CALL software preferences (see FID measure in the Appendix). The measure was constructed and its content analyzed on the basis of hypotheses about how the FID construct—as defined by Witkin and Goodenough (1981)—would affect learners’ behavior during the use of a sophisticated tutorial CALL program such as E-Tutor.

Procedure
During the 15 weeks of system use in spring semester 2006, the E-Tutor logged the student-computer interactions including students’ use of help aids. Student use of the system was substantial enough that the log data provided ample evidence of their working styles. During the 12th week of the semester, students responded to the 30-items on the FID-CALL questionnaire.

Data analysis began by estimating the reliability of the FID-CALL questionnaire and identifying questions for which respondents were in nearly complete agreement. To allow for cross-comparison among learners, we normalized the working style data for all students by dividing total number of times the students accessed help in the program (grammar aids and dictionary) by the number of vocabulary and grammar errors, respectively, to obtain a measure of rate of help use per error commission. We calculated the number of skips and peaks divided by the total number of submissions for each student to obtain comparable measures across students. We then examined the records of learners’ behavior to identify individual variation. We compared selected students’ responses from the questionnaire with the data obtained from the E-Tutor logs, focusing on five questions from the questionnaire that make explicit reference to computer help aids. Accordingly, we excluded questions that do not specifically indicate whether the help aid is part of the CALL program or whether it is provided in print form and therefore external to the software (e.g., a print dictionary). Moreover, we only considered questions that contrast learners using computer aids versus solving the task by themselves. Thus, we ignored questions that contrast students using computer aids as opposed to those that elicit help from an instructor.
RESULTS
The results indicated that the FID-CALL survey failed to find reliable individual differences among the participants in the study, but the examination of individual items revealed some areas of agreement on CALL design among students. Individual variation was found in the records of students’ working-style behaviors, and significant differences were linked to corresponding questionnaire responses.

Reliability
Internal consistency reliability refers to the extent to which a scale consistently distinguishes among participants. Measurement specialists would agree that researchers are not warranted in making inferences about the variation in characteristics among individuals unless the scores obtained from the scale are estimated to have a high degree of reliability. The estimated Cronbach’s alpha reliability was .27 for the 30-item scale. This indicates that the scale as a whole did not reliably distinguish among respondents, and therefore the scores from the complete scale were not used for subsequent analysis. Instead, we examined responses from individual questions, which provided some insights into students’ opinions of CALL design and links to behaviors. These may be useful for subsequent development of a more robust scale.

Unanimous Opinions
The results displayed in Table 1 indicate that students unanimously agree (GE3 mean = .00) that they prefer working with CALL programs that include help aids (e.g., transcripts, dictionary, and grammatical paradigms). Moreover, for questions that relate to dictionary use, students generally agree that they choose to look up a word in the dictionary (V1 mean = .98, R3 mean = .82). Interestingly, there is also strong agreement among our learners that, when working with software intended for language learning, students commonly prefer working alone as opposed to working with a human tutor (GE5 mean = .82). Finally, with respect to grammar instruction, students generally prefer graded exercises that proceed from easy to difficult (G3 mean = .04) as well as exercises that relate new content to previously learned material (G6 mean = .06).

Table 1
Descriptive Statistics for the Questionnaire Responses (N = 50)

<table>
<thead>
<tr>
<th>Question</th>
<th>Item Mean (High = FI; Low = FD)</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>GE3: You prefer working with software intended for language learning that a) includes help aids (e.g., transcripts, dictionary, etc.), b) has no help aids.</td>
<td>.00</td>
<td>.000</td>
</tr>
<tr>
<td>G3: When working with grammar exercises, it is most likely that you would a) start working with easy grammatical exercises, b) start working with more challenging grammatical exercises.</td>
<td>.04</td>
<td>.198</td>
</tr>
<tr>
<td>G6: It is easier for you to remember how to use a grammatical structure if a) it doesn’t relate to grammar structures previously studied, b) it relates to structures previously studied.</td>
<td>.06</td>
<td>.240</td>
</tr>
</tbody>
</table>
L4: After listening to an online lecture, it is most likely that you will remember a) concepts explained through examples, b) concepts explained by definition. .12 .328

G5: If the computer shows that a sentence you wrote is grammatically incorrect, it is most likely that you would a) use the grammar check tool, b) not use the grammar check tool. .14 .351

GE4: You would prefer working with software intended for language learning that includes a) topics you are familiar with, b) topics that are new to you. .20 .404

R5: When reading an online text for pleasure, it is most likely that you will a) focus on the specific details, b) try to get the general idea of the text. .20 .404

R1: When reading online texts, you prefer a) readings on topics that are familiar to you, b) readings on topics that are new to you. .24 .431

W3: When you are writing a report based on a lecture and the computer indicates that you have made a grammar or spelling mistake, it is most likely that you would a) use some of the help options provided by the computer, b) try to correct the mistake on your own. .30 .463

R2: In a reading assignment for a class, it is most likely that you would read a) what you have been asked to read even if you know about the topic, b) what you consider you need to read. .30 .463

V5: You are more likely to remember the meaning of a word if a) you read the definition of the word, b) you read a phrase where the word is used. .30 .463

L3: When listening to lectures, you prefer a) lectures on topics you are familiar with, b) lectures on topics that are new to you. .34 .479

L2: In class, if you don’t understand a listening comprehension exercise you most likely would a) ask your classmate/teacher to explain it, b) check the transcripts. .34 .479

GE1: Using computers to learn a language seems more attractive to you if a) you decide on the type of exercises you want to work on, b) the computer guides you and suggests what exercises to do. .34 .479

W1: When taking notes based on a lecture you most likely would a) write down your own ideas about the content of the lecture, b) write down excerpts of the content of the lecture. .38 .490

L1: If you do not understand a part of an online lecture, you most likely would a) read the transcripts while you listen to the lecture, b) listen to the lecture again, before using the transcripts. .38 .490

G4: If after receiving computer-generated feedback you discover that a grammar answer you entered is incorrect you most likely would a) try as many times as necessary until you get it right, b) try one or two more times and then use the check answer option. .42 .499

R4: After reading an online text for a class assignment, it is most likely that you would a) talk about it with a classmate to clarify some ideas b) trust your own understanding of the text. .44 .501
Individual Differences in Behavior

The unit of analysis for the working style data was the sentence. The participants completed three different types of exercises: a dictation (transcribing a German story or a dialogue), a translation (from English into German), and a sentence-building task (forming sentences from uninflected base forms and inserting missing function words).
Over the course of the semester, we collected a total of 46,025 sentences. Without any data normalization at this point, this resulted in an average of 920 sentences per student. Moreover, on average students skipped 18 exercises (6 students did not skip any exercises), they looked up an answer 53 times (all students peeked at some point), and made an average of 258 mistakes (12,905 mistakes in total). On average, students also accessed the dictionary 97 times and the grammar aids 27 times.

The computer logs of the 50 study participants further indicate that

1. 3 out of 50 students did not access the grammar aids (dynamic inflectional paradigms, static notes on word order, etc.) but did access vocabulary in the dictionary,
2. 7 out of 50 students did not access the dictionary but did access grammar aids, and
3. 2 out of 50 students accessed neither the grammar aids nor the dictionary.

In other words, despite the overwhelming positive responses about help that students expressed on the questionnaire, 12 students did not make use of one or both types of help. While one might expect to find a connection between students’ access of help aids and the number of errors they made when completing exercises, further analysis of our data did not confirm this expectation.

**Relationship Between FID Questions and Behavior**

The five questions that we chose from the questionnaire pertain to the use of help options and showed some individual difference in students’ responses. We examined the behavior associated with responses on five questions, and having found a trend that made sense, we combined the five questions into a single scale for which we calculated reliability and correlations between the scores and the behaviors of interest.

**Analysis of individual items**

For a detailed analysis of student responses in the questionnaire and the E-Tutor log, we applied an independent samples t test by grouping students according to the answers they provided for each question in the questionnaire (1 or 0). The dependent variables were quantitative data from students’ access of help in the program, which, as described above, were calculated for each student by dividing the number of times the student accessed the dictionary and grammar aids by the number of vocabulary and grammar errors, respectively, to obtain a measure of rate of help use per error commission. We also calculated the number of skips and peeks divided by the total number of submissions for each student to obtain measures that were normalized in view of the different numbers of sentences that the students worked on.

The four types of help aids we consider for this analysis provide a range of learner assistance for solving a task. At one end of the spectrum, the *E-Tutor* provides the possibility for the student to access the correct answer (peek) for an exercise item. This is the most explicit form of help in that, instead of working with hints to derive a correct answer, the student simply clicks on a button to see the correct answer. On the other end of the spectrum, skipping
an exercise (skip) provides the least amount of help. Here, the software provides the option for the student to terminate and abandon the task without participating in the error correction process. Accordingly, this button provides no help whatsoever to solve the current task. The online dictionary and the grammar aids provide additional information useful for solving the task but, nonetheless, requires the student to do some thinking to derive the correct solution, more so in the use of grammar aids than in the dictionary.

Tables 2-6 display the results of the statistical analyses for the questions considered in the questionnaire.

Table 2
Comparison of Mean Help Use Frequencies for Students Responding Positively and Those Responding Negatively to Grammar Question 1

<table>
<thead>
<tr>
<th></th>
<th>Yes (FD) (n = 24)</th>
<th>No (FI) (n = 26)</th>
<th>t(48)</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Skip</td>
<td>Mean: 0.018, SD: 0.026</td>
<td>Mean: 0.018, SD: 0.029</td>
<td>0.040</td>
<td>.770</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dictionary</td>
<td>1.671, 2.323</td>
<td>0.996, 2.218</td>
<td>-1.051</td>
<td>.312</td>
</tr>
<tr>
<td>Grammar aid</td>
<td>0.143, 0.112</td>
<td>0.076, 0.074</td>
<td>-2.490</td>
<td>.018</td>
</tr>
<tr>
<td>Peek</td>
<td>0.066, 0.059</td>
<td>0.043, 0.033</td>
<td>-1.671</td>
<td>.006</td>
</tr>
</tbody>
</table>

Table 2 shows that, given the choice, slightly fewer students (24) would choose computer-based help aids in this situation. Moreover, the students who indicated in the questionnaire that they most likely would use help options (the FD option) indeed accessed the grammar aids of the E-Tutor significantly more often than the FI students (t(48) = -2.490, p < .05). They also looked up answers significantly more often (t(48) = -1.671, p < .01). If students suspect that a grammar-based answer is wrong, they most likely would use the help option to figure out what the mistake might be before checking it (FD), as opposed to trying again without using any help from the computer (FI).

Table 3
Comparison of Mean Help Use Frequencies for Students Responding Positively and Those Responding Negatively to Grammar Question 4

<table>
<thead>
<tr>
<th></th>
<th>Yes (FD) (n = 29)</th>
<th>No (FI) (n = 21)</th>
<th>t(48)</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Skip</td>
<td>Mean: 0.017, SD: 0.024</td>
<td>Mean: 0.020, SD: 0.031</td>
<td>0.378</td>
<td>.108</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dictionary</td>
<td>1.612, 2.837</td>
<td>0.916, 1.039</td>
<td>-1.070</td>
<td>.018</td>
</tr>
<tr>
<td>Grammar aid</td>
<td>0.119, 0.116</td>
<td>0.092, 0.069</td>
<td>-0.946</td>
<td>.002</td>
</tr>
<tr>
<td>Peek</td>
<td>0.064, 0.052</td>
<td>0.041, 0.041</td>
<td>-1.648</td>
<td>.125</td>
</tr>
</tbody>
</table>

For question 4 of the grammar section, students who indicated that they would more likely rely on computer aids (the FD response) indeed accessed both the dictionary and the grammar aids significantly more often (t(48) = -1.070, p < .05; t(48) = -0.946, p < .005). If after receiving computer-generated feedback students discover that a grammar answer they entered is incorrect, they most likely would try one or more times and then use the check answer button (FD), as opposed to trying as many times as necessary until they get it right (FI).
Table 4
Comparison of Mean Help Use Frequencies for Students Responding Positively and Those Responding Negatively to Grammar Question 5

<table>
<thead>
<tr>
<th></th>
<th>Yes (FD) ((n = 43))</th>
<th>No (FI) ((n = 7))</th>
<th>(t(48))</th>
<th>(p)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Skip</td>
<td>0.020 0.028</td>
<td>0.004 0.004</td>
<td>-1.456</td>
<td>.018</td>
</tr>
<tr>
<td>Dictionary</td>
<td>1.448 2.421</td>
<td>0.529 0.513</td>
<td>-0.993</td>
<td>.078</td>
</tr>
<tr>
<td>Grammar aid</td>
<td>0.117 0.102</td>
<td>0.053 0.055</td>
<td>-1.609</td>
<td>.047</td>
</tr>
<tr>
<td>Peek</td>
<td>0.056 0.048</td>
<td>0.040 0.048</td>
<td>-0.813</td>
<td>.861</td>
</tr>
</tbody>
</table>

Similar to the previous question, students who indicated that they would use the grammar check tool on question 5 of the grammar section (the FD response) accessed the grammar aids and skipped exercises significantly more often \((t(48) = -1.609, p < .05; t(48) = -1.456, p < .05)\). If the computer shows that a sentence students wrote is grammatically incorrect, it is most likely that they would use the grammar check tool (FD), as opposed to not using the grammar check tool (FI).

Table 5
Comparison of Mean Help Use Frequencies for Students Responding Positively and Those Responding Negatively to Grammar Question 3

<table>
<thead>
<tr>
<th></th>
<th>Yes (FD) ((n = 35))</th>
<th>No (FI) ((n = 15))</th>
<th>(t(48))</th>
<th>(p)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Skip</td>
<td>0.018 0.024</td>
<td>0.020 0.034</td>
<td>0.251</td>
<td>.124</td>
</tr>
<tr>
<td>Dictionary</td>
<td>1.538 2.538</td>
<td>0.810 0.142</td>
<td>-1.039</td>
<td>.171</td>
</tr>
<tr>
<td>Grammar aid</td>
<td>0.122 0.111</td>
<td>0.075 0.052</td>
<td>-1.541</td>
<td>.001</td>
</tr>
<tr>
<td>Peek</td>
<td>0.064 0.053</td>
<td>0.031 0.024</td>
<td>-2.251</td>
<td>.013</td>
</tr>
</tbody>
</table>

For question 3 of the writing section, students who indicated that they would use help aids in case of a grammar or spelling mistake (the FD response), accessed the grammar aids and looked up a correct answer significantly more often \((t(48) = -1.541, p < .005; t(48) = -2.251, p < .05)\). When students are writing a report based on a lecture and the computer indicates that they have made a grammar or spelling mistake, it is most likely that they would use some of the help options provided by the computer (FD), as opposed to trying to correct the mistake by themselves (FI).

Table 6
Comparison of Mean Help Use Frequencies for Students Responding Positively and Those Responding Negatively to Grammar Question 2

<table>
<thead>
<tr>
<th></th>
<th>Yes (FD) ((n = 26))</th>
<th>No (FI) ((n = 24))</th>
<th>(t(48))</th>
<th>(p)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Skip</td>
<td>0.017 0.032</td>
<td>0.019 0.021</td>
<td>-0.244</td>
<td>.259</td>
</tr>
<tr>
<td>Dictionary</td>
<td>1.522 2.323</td>
<td>1.101 2.243</td>
<td>0.650</td>
<td>.703</td>
</tr>
<tr>
<td>Grammar aid</td>
<td>0.127 0.114</td>
<td>0.091 0.082</td>
<td>-1.302</td>
<td>.057</td>
</tr>
<tr>
<td>Peek</td>
<td>0.058 0.058</td>
<td>0.049 0.036</td>
<td>0.668</td>
<td>.025</td>
</tr>
</tbody>
</table>
The data displayed in Table 6 indicate that students who prefer programs that allow them to skip exercises look up answers significantly more often ($t(48) = 0.668$, $p < .05$), possibly implying that if learners choose not to solve a task, they are, nonetheless, curious about the correct solution. Along the same lines and, although not statistically significant, these students generally seem more likely to make use of other help aids. It is interesting, however, that the students who indicated that they prefer programs that allow them to skip exercises actually skipped fewer exercises than the other group of learners. Students prefer working with software intended for language learning that allows them to skip exercises (FI), as opposed to requiring them to complete exercises (FD).

Table 7 provides a summary of the comparisons resulting in significant differences. To adjust for the multiple $t$ tests for each type of help aid, the table below shows those results with a significance level of less than .025 (.05 x 5). When students are divided by the FI and FD responses on the first four questions referring to explicit use of computer aids, significant differences were found in their use of several types of help. The question asking about students’ preferences for skipping exercises resulted in differences among students, but no significant differences in behavior among the FI and FD responses. These findings are consistent with the perspective of the FD student as the one who relies on external help even if the help comes from the computer.

Table 7
Summary of Significant Differences ($p < .025$) in Behaviors between FI and FD Responders.

<table>
<thead>
<tr>
<th>Question</th>
<th>Skip</th>
<th>Dictionary access</th>
<th>Grammar aid access</th>
<th>Peek</th>
</tr>
</thead>
<tbody>
<tr>
<td>G1: Use help immediately if you don’t know the answer</td>
<td>.018 FD</td>
<td>.006 FD</td>
<td></td>
<td></td>
</tr>
<tr>
<td>G4: Use the check answer option after trying and receiving feedback about incorrectness of response</td>
<td>.018 FD</td>
<td>.002 FD</td>
<td></td>
<td></td>
</tr>
<tr>
<td>G5: Use grammar help to get error explanation</td>
<td>.018 FD</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>W3: Use grammar help to receive help with marked error</td>
<td>.001 FD</td>
<td>.013 FD</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GE2: Like to be able to skip exercises</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Analysis with the Five-item Scale**

The results obtained from the analysis of the individual items were sufficiently consistent to warrant exploration of those items as constituting a scale on its own. A five-item scale would not be expected to have high reliability, and therefore the estimated Cronbach’s alpha of .40 suggests some promise for the use of items of this type to identify individual differences, particularly when compared to the Cronbach’s alpha of .27 for the entire scale. The fact that this reliability is considerably higher than that for the entire 30-item scale suggests that these five items might provide some useful insight into how to assess FID in CALL. Having obtained a score with some reliability, we looked at the association of the scale score with the participants’ behaviors.
As Table 8 shows, the trends that appeared in the item-level analysis were also evident when the five-item scale was used. The negative correlations between the five-item scale and the use of the help aids suggests, as predicted, that the FI students were the ones who tended not to use the help. The correlations are statistically significant in the case of the use of grammar aids and peeking at the correct answers. However, in view of the low reliability of the five-item questionnaire, the more meaningful correlations are the disattenuated ones that appear below the observed correlations. These are the magnitude of correlations that one would expect to find if the scale were completely reliable, and therefore the content offers a suggestion for development of a new FID-CALL scale. A scale with additional items referring to computer help options would be expected to produce scores with greater reliability when used with appropriate students. Moreover, a Likert-item response consisting of a scale such as "always, frequently, sometimes, never" would provide greater variance in item responses and therefore hold the potential for scores with increased reliability.

<table>
<thead>
<tr>
<th></th>
<th>Skip</th>
<th>Dictionary access</th>
<th>Grammar aid access</th>
<th>Peek</th>
</tr>
</thead>
<tbody>
<tr>
<td>Observed Pearson Correlation</td>
<td>-.034</td>
<td>-.176</td>
<td>-.404</td>
<td>-.292</td>
</tr>
<tr>
<td>(Disattenuated Correlations)*</td>
<td>(-.054)</td>
<td>(-.280)</td>
<td>(-.642)</td>
<td>(-.464)</td>
</tr>
<tr>
<td>(p)</td>
<td>.815</td>
<td>.666</td>
<td>.004</td>
<td>.040</td>
</tr>
</tbody>
</table>

*Calculated on the basis of the estimated Cronbach’s alpha reliability of .40 for the five-item scale and assumed perfect reliability of the working style data.

**DISCUSSION**

The results of the project indicated that the FID-CALL survey we investigated (Cárdenas-Claros, 2005) did not reliably distinguish among participants in this study. These data provide additional evidence that more work is needed to develop a context-specific measure of FID for CALL, and the analysis pointed to some useful findings for such an endeavor. First, we identified several questions on which virtually no individual differences were detected. These questions are of interest from the perspective of CALL design (e.g., all students want to have help options available) and from the perspective of future questionnaire development. Questions that fail to detect differences among learners are not useful for a survey aiming to detect individual differences.

Second, we identified some survey questions that were particularly good at distinguishing individual differences, and we showed that these questions predicted individual differences in behavior pertaining to the use of help as well. The analysis of the behavioral data by students’ responses to five individual questions revealed the hypothesized trend of an association between the FD response and the use of help. Having found this trend, or the appearance of consistent item functioning with respect to prediction of behavior, we looked at what the reliability of such a scale would be. The Cronbach’s alpha of .40 for the five-item scale suggests that these questions are headed in the right direction in the pursuit of a reliable measure. Such a measure would need to be developed by adding additional items that
target the learners’ preferences for relying on the computer for help versus counting on their own knowledge and persistence.

Some support for the validity of inferring that scores on such a scale are indicative of FID in CALL use comes from correlations between the five-item FID-CALL scale and learners’ actual use of various levels of help. Figure 2 places the types of help on a continuum from those providing the least guidance to those providing the most guidance.

Figure 2
Disattenuated Correlations Between Five-item FID-CALL Scale and Observed Behaviors Shown on a Continuum from Those Requesting Least to Most Guidance

![Figure 2](image)

The correlations shown in this figure are disattenuated, that is, those that would be expected with these data if the FID-CALL scale provided completely reliable scores. In other words, if the five-item FID-CALL scale were expanded to produce reliable scores, we would expect the correlations with actual behavior to be of the magnitude shown in Figure 2. Selection of the two forms of help providing the most guidance would produce moderate correlations with the FID-CALL scale, whereas the one providing the least guidance would not correlate at all. The negative correlations between the five-item FID-CALL scale indicate that the FD learners tend to use the help when they make errors. This fits the theoretical prediction one would make on the basis of the FID construct, based on the assumption that computer help is viewed by the learner as an external aid, just as help from another person.

If these findings can be replicated with a better developed FID-CALL survey, the results can be interpreted as shedding light on the theoretical definition of FID in CALL. In the five items of the questionnaire that we examined, the assumption was that the learner sees the computer as taking on a similar role as an instructor or peer. The FI student then would not feel inclined to consult the available computer-assisted help (A in Figure 3), whereas the FD student would use the computer help like that provided by other students or the teacher. Another potential view of the computer is a device for extending the self reliant FI students’ reach in performing individually (B in Figure 3), and the FD student not seeing it as the type of human help that he or she seeks. View B, however, is not supported by these data.
CONCLUSION

This research has made some progress toward the goal of better understanding the intersection of individual differences with language learners’ use of CALL. The FID-CALL survey failed to find reliable individual differences among the participants in the study, but examination of performance on individual items along with learners’ behavior over the course of a 15-week semester provided a means for proposing some steps forward. In particular, some areas of agreement (i.e., lack of individual variation) on CALL design among students were identified. Individual variation was found in records of students’ working style behaviors, and significant differences were linked to corresponding questionnaire responses. Specifically, the five questions chosen from the questionnaire are clear areas of individual differences in view of student responses. The most significant differences were found with students’ access of grammar aids and correct answers. In almost all of the cases where significant differences were found, the FD responses on the questions were associated with greater use of help. This finding is consistent with the perspective of the FD student as the one who relies on external help even if the help comes from the computer.
This research has also revealed the complexity of the study of individual differences in a manner that is relevant to CALL. Assessment of IDs cannot be accomplished by an off-the-shelf measure that has been “validated” by experts and can, therefore, be assumed relevant. Instead, researchers in assessment agree that the inferences and uses of any assessment have to be validated by users. One important step in this process, of course, is to demonstrate the reliability of the measure for the sample about which one wishes to infer individual differences. However, the critical stage prior to developing a survey for assessment of an individual difference is to develop the theoretical definition of the construct in a manner relevant to the intended explanation. In CALL research, such context-specific construct definition explores new ground as trait-type constructs developed in the past are delimited through the learning opportunities provided by CALL. Despite this complexity, however, it seems that the need exists to better understand how CALL intersects with individual differences—whether they be differences in FID or in other constructs of individual differences.

NOTE
1 A previous study by Heift (2006), for instance, showed that learner access of context-sensitive help differs significantly with respect to feedback, exercise type, and language proficiency.

REFERENCES


APPENDIX

Questionnaire

I. General (GE)

1. Using computers to learn a language seems more attractive to you if
   a) you decide on the type of exercises you want to work on.
   b) the computer guides you and suggests what exercises to do.

2. You prefer working with software intended for language learning that
   a) requires you to complete all the exercises.
   b) allows you to skip some exercises.

3. You prefer working with software intended for language learning that
   a) includes help aids (e.g., transcripts, dictionary, etc.)
   b) has no help aids.

4. You would prefer working with software intended for language learning that includes
   a) topics you are familiar with.
   b) topics that are new to you.

5. When working with software intended for language learning you prefer
   a) working alone.
   b) working with a human tutor.

II. Listening (L)

1. If you do not understand a part of an online lecture, you most likely would
   a) read the transcripts while you listen to the lecture.
   b) listen to the lecture again, before using the transcripts.

2. In class, if you don’t understand a listening comprehension exercise you most likely would
   a) ask your classmate/teacher to explain it.
   b) check the transcripts.

3. When listening to lectures, you prefer
   a) lectures on topics you are familiar with.
   b) lectures on topics that are new to you.

4. After listening to an online lecture, it is most likely that you will remember
   a) concepts explained through examples.
   b) concepts explained by definition.

III. Reading (R)

1. When reading online texts, you prefer
   a) readings on topics that are familiar to you.
   b) readings on topics that are new to you.
2. In a reading assignment for a class, it is most likely that you would read
   a) what you have been asked to read even if you know about the topic.
   b) what you consider you need to read.

3. When reading a text on a topic you are not familiar with, it is most likely that you would
   a) use the dictionary to better understand the text.
   b) not use the dictionary.

4. After reading an online text for a class assignment, it is most likely that you would
   a) talk about it with a classmate to clarify some ideas.
   b) trust your own understanding of the text.

5. When reading an online text for pleasure, it is most likely that you will
   a) focus on the specific details.
   b) try to get the general idea of the text.

IV. Writing (W)

1. When taking notes based on a lecture you most likely would
   a) write down your own ideas about the content of the lecture.
   b) write down excerpts of the content of the lecture.

2. When writing a report based on a lecture, it is most likely that you would
   a) highlight your opinion about the content of the lecture.
   b) provide a description of the content of the lecture.

3. When you are writing a report based on a lecture and the computer indicates that you
   have made a grammar or spelling mistake, it is most likely that you would
   a) use some of the help options provided by the computer.
   b) try to correct the mistake on your own.

4. If you don’t know how to spell a word when writing a report based on a lecture, you most
   likely would
   a) ask the person besides you to spell it for you.
   b) look it up in the dictionary.

V. Vocabulary (V)

1. If you want to know the meaning of a new word, you most likely would
   a) look it up in the dictionary.
   b) not look it up in the dictionary.

2. If you don’t know the meaning of a word, you most likely would
   a) ask a classmate or the instructor for the meaning.
   b) look it up in the dictionary.

3. If you encounter a word that you are not familiar with, you most likely would
   a) stop reading and look it up in the dictionary or in any other help provided by the
      computer.
   b) keep on reading and get the meaning from the context.
4. After learning a new word, you most likely would
   a) try to use the word as much as possible.
   b) not use the word unless you are completely sure when to use it.

5. You are more likely to remember the meaning of a word if
   a) you read the definition of the word.
   b) you read a phrase where the word is used.

VI. Grammar (G)

1. If you suspect that a grammar-based answer is wrong, you most likely would
   a) use the help option to figure out what the mistake might be before checking it.
   b) try again without using any help from the computer.

2. During a writing class where you are working with computers, if you are not sure a sen-
   tence is grammatically correct, it is most likely that you would
   a) ask a classmate to check it for you.
   b) use the grammar checking tool provided by the computer program.

3. When working with grammar exercises, it is most likely that you would
   a) start working with easy grammatical exercises.
   b) start working with more challenging grammatical exercises.

4. If after receiving computer-generated feedback you discover that a grammar answer you
   entered is incorrect you most likely would
   a) try as many times as necessary until you get it right.
   b) try one or two more times and then use the check answer option.

5. If the computer shows that a sentence you wrote is grammatically incorrect, it is most
   likely that you would
   a) use the grammar check tool.
   b) not use the grammar check tool.

6. It is easier for you to remember how to use a grammatical structure if
   a) it doesn’t relate to grammar structures previously studied.
   b) it relates to structures previously studied.

7. When learning a new grammatical structure, you are most likely to
   a) try to use this structure even if you are not sure if it is right.
   b) avoid using this structure until you feel comfortable using it.
AUTHORS’ BIODATA

Carol A. Chapelle (Ph.D. in Second Language Acquisition and Teacher Education, University of Illinois) is Professor of TESL/Applied Linguistics at Iowa State University. She teaches classes in CALL, second language testing, and second language acquisition in an MA program in TESL/Applied Linguistics with specialization in CALL and in a Ph.D. program in Applied Linguistics & Technology. Carol has published widely in these areas. She has served as editor of TESOL Quarterly and president of the American Association for Applied Linguistics.

Dr. Trude Heift is Associate Professor of Linguistics at Simon Fraser University, Vancouver, Canada. Her research focuses on ICALL, particularly, on intelligent tutoring systems, human-computer interaction, and student modeling. She is the author of E-Tutor, an NLP-based online language-learning tool for German.

AUTHORS’ ADDRESSES

Carol Chapelle
TESL/Applied Linguistics
Iowa State University
Email: carolc@iastate.edu

Trude Heift
Linguistics Department
Simon Fraser University
Email: heift@sfu.ca