SECOND LANGUAGE STUDENTS' CONFIDENCE IN PRODUCING THE CORRECT ANSWER

Seime Kebede*

Abstract

This paper is designed to assess whether there was any relationship between correctness in answering and accuracy in judging the correctness / incorrectness of the answers the students wrote. To achieve this aim, a test was set and administered to 29 students studying Physics in the Kiremt Programme at Bahir Dar Teachers' College. Every student was required to indicate how sure he was of the correctness of the answer he / she gave to each item by writing one number of the 5 point scale.

The results indicate that the subjects neither rated 5 each of the items they got correct nor 1 all those items they were wrong in. The students' confidence ratings are discussed in terms of Monitor Hypothesis, fossilization, and teaching.

---

*Address: Bahir Dar Teachers' College, P.O. Box 76, Bahir Dar, Gojam.

Introduction

Scholarly interest in the relation between learner characteristics and second language learning is not new. Early studies took into account cognitive variables such as intelligence and aptitude (Caroll 1962; Pimsleure 1966); later research focused on such affective variables as attitude and motivation (e.g. Lambert 1972; Oller et al 1977). More recently, affective variables such as confidence in listening tasks have been examined by Yule eta al (1985). It is not the intention of this study to present an extensive review of this literature but rather to underline the fact that over time, this body of research has contributed much to the teaching and learning process in general, and to the teaching of second language in particular.

The purpose of the study

This paper attempts to study whether there is any relationship between correctness in answering and accuracy in judging the correctness/incorrectness of the answers the students wrote. More specifically, the study is designed to:

a) ascertain the validity of the assumption that a learner's accuracy in answering a test item reflects knowledge and inaccuracy lack of knowledge of a linguistic element.

b) investigate the possibility that some learners choose correct answers with no confidence whatsoever in their accuracy and choose wrong answers with a lot of confidence in their accuracy.

Subjects and Method

Subjects:- An experiment was conducted which was designed to assess adult second language learners' production confidence in some selected grammatical elements. The experiment was administered to 29 first year diploma students studying physics in the Kiremt Programme at Bahir Dar Teachers' College.

At the time the research was carried out, the sample subjects were offered a 48 hr. course in English and other 128 hr. courses in the medium of English at the government's expense.

The major criteria for entering the programme are both objective and subjective: (a) ESLCE, (b) Service in the teaching profession, (c) area of specialization and (d) Social Services rendered to the community outside school hrs. Maximum weight is placed on measures (a) and (b).

Although the combination of subjective and objective entrance criteria make it difficult to strictly compare subjects on the objective measures, the two achievement test results in Freshman English course which were administered to this same group prior to the experiment, show that there is no marked difference in their proficiency level of the language.

In order to enhance the likelihood that the findings of this study would be reliable and not due to chance, the researcher sampled all the class. The students were tested in the final weeks of their two months programme.

Method

The production task consisted of five grammatical features: phrasal verbs, prepositions, quantifiers, tenses, and reported speech. The test contained a total of 27 items. For each item the subjects had to choose or supply the answer by themselves.

After giving the answers the students were required to indicate how sure they were of the correctness of their answer by writing one number from among the 5 point scale on the space provided. That is, the students were asked to rate their 'confidence' in the correctness of their answers using a five-point Likert scale.

The scale:- 5 completely sure
4 above average sure
3 average "
2 below average "
1 not at all "

The students were given orientation on how to use the 5 point scale. Nevertheless, some response effects (such as, self - presentation, response set, acquiessence, etc) are likely to surface when using a five point scale in a questionnaire design.

Results and Analysis:- The learners performance on each grammatical element and their confidence ratings that is, 'confidence' in the correctness of their answers are illustrated in the table below (cf. Table 2. Appendix B, and C).

Table 1. Correct/incorrect mean confidence and % score per grammatical unit.

Grammatical units	Mean confidence ratings for		% Score	
	Correct answers	Wrong answers	Correct answers	Wrong answers
Phrasal verbs	4.22	3.18	54	46
Prepositions	4.58	3.77	70	30
Quantifiers	4.60	4.33	94	6
Tenses	3.53	3.72	44	56
Reported speech	4.74	3.92	57	43
T O T A L	21.67	18.52	320	180
Grand Mean	4.33	3.78	64	36

As could be noted from the above table, the students, though with a varying degree, are above average (cf. The scale p.2) in their confidence ratings of their actually correct answers. In fact they are more sure of their knowledge in reported speech, quantifiers, prepositions and phrasal verbs.

Conversely, the students are not aware of their knowledge in tenses. Regarding this grammatical unit, the students' mean confidence for actually wrong answers (3.72) is greater than for actually right answers (3.53). Their 'inability' in this aspect of grammar is further proved by the score they obtained (cf. Table 2). It, thus, appears that the students are having trouble with this particular element of grammar. However, they tend to show that they are 'sure' in producing the correct answers when in fact they are wrong. Since they are confident in wrong answering, they are not inhibited by the fact that they are making mistakes when using the tense of the language. These students who are confident in wrong answering and are not aware of their mistakes when using the tense of the language seem to have the tendency to be outgoing and interested in people and things in the environment (extraversion). According to Stern (1983: 379), "it is sometimes said that outgoing students with histrionic talents are more successful language learners than more inhibited or introverted students."

In contrast to the other grammatical elements tested, with the exception of the reported speech, tenses are full of "meaning potential" and can be taught with lots of meta-explanations. Quantifiers, prepositions and phrasal verbs are favourites for uncontextualized language exercises and tests. Besides, teachers tell their students that these areas are difficult. Consequently, they are simply being memorized but not worried about or adjusted as they are peripheral and low in meaning potential. Thus, these grammatical areas seldom lead to confusion when deciding the correct form of the language. On the other hand, because they have lots of meta-explanations, tenses require a developed monitor to be used that often lead to hesitation as to which form is 'correct'. Perhaps this knowledge of 'inability' that is low confidence in tenses comes from 'over' teaching or meta - explanation and therefore the development of the monitor.

Again, though to a lesser extent, the students' confidence ratings of their ability to use quantifiers is misleading. There isn't as such marked difference in their ratings between the correct and the wrong answers. This demonstrates that the students have a wrong conception of their ability in quantifiers. Specifically, regarding this particular grammatical unit, about 43% of the students indicated above-average confidence when in fact they produce wrong answers (cf. Table 2). If the students were sure of the answers they gave, they should have indicated low confidence in their wrong answers since they obtained a mean score of 93% for their correct answers. The likely assumption here is that the students made use of their partial knowledge, and the chance factor (50% cf. Appendix A & D) and took risk in answering the question.

In other words, multiple choice test can be viewed as a decision making-task because of the possibility that examinees could enhance their score by guessing. Thus, guessing when confidence exceeds the level of total doubt (that is complete uncertainty) will yield performances in excess of chance. There can be little doubt that highly confident students might use partial information more effectively than do their more cautious peers. The less amount of knowledge that a subject possess, the higher he would hesitate to risk in his response to the multiple choice item. There seems to exist a significant relation among strong risk - taking despositions and higher mean confidence, and score for correct answer.

The students had a good chance of being correct, because (apart from 50% guessing factor) they possess enough information to make some headway towards the elimination of the incorrect alternative as this test was looking for the learned rather than the acquired system. If they make good use of partial information this kind of guessing will enhance their score.

In addition as there was no penalty for wrong answers the students employ a strategy indicating a wish to maximize the number of right answers which in effect would be greater mean score. Perhaps, these could be the reasons why the two mean confidence ratings (correct and incorrect) are almost equal, and the mean score for correct answer higher on the test of quantifiers.

In contrast, the test on preposition and reported speech were not multiple choice, and hence the students were not only deprived of the chance of blind guessing but also the possibility of making some headway towards the elimination of incorrect alternatives. As a consequence it seems plausible to hypotesize that the students rated their confidence ratings for wrong answers lower and obtained a relatively higher score for the correct answers.

In a test result such as these, in which students are not sure/aware of their ability, and are reluctant to take risk, the teacher's task would be to rectify these items that the students gave wrong answer to. But in a test result such as quantifiers, it is evident that the teacher should review not only the wrong answers but also those items which the students were unsure about, yet recorded correctly.

Regarding the students' performance on individual items the table below illustrates their results.

Table 2 Results of the students on individual items.

Item	No. of students who got the item correct	X confidence for correct answers	No. of student who got the item wrong	X confidence for wrong
1	8	4.25	21	3.81
2	14	4.29	15	3.31
3	20	4.35	9	2.56
4	17	4.35	12	3.17
5	20	3.90	9	2.67
6	23	4.56	6	4.00
7	23	4.56	6	3.33
8	19	4.31	10	3.5
9	20	4.60	9	3.67
10	24	4.50	5	3.60
11	26	4.85	3	3.67
12	27	4.96	2	1.00
13	7	4.00	22	4.68
14	15	4.67	14	3.14
15	20	4.30	9	3.78
16	28	4.82	1	5.00
17	25	4.28	4	3.50
18	25	4.84	4	5.00
19	29	4.51	-	-
20	29	4.55	-	-
21	23	3.96	6	3.83
22	7	3.14	22	3.27
23	9	4.00	20	3.75
24	8	3.88	21	3.48
25	24	4.62	5	4.60
26	15	4.93	14	4.57
27	11	4.72	18	3.50

As could be observed from Table 2, while indicating their confidence, the students neither rated all their correct answers "5" nor all wrong answers "1". Their overall mean confidence ratings for an individual incorrectly answered item ranges from 1 to 5 while for correct answers it is between 3.14 - 4.96. The average confidence rating for correct answers ($\bar{X}$: 4.37) is higher than the average for incorrect answers (3.78) (cf. Table 2 and Appendix B.) This shows that the learners are more sure about their correct answers than their incorrect ones. In other words, as opposed to the wrong answers, the students indicated a higher mean confidence ratings for correct answers with the exceptions of one item on prepositions (item number 13), two items on Quantifiers (items number 16 and 18), and one item on tense (item number 22). There is also a case where the mean confidence is equal for both correct and incorrect items (items 25 on reported speech). ref. Table 2.

On the other hand, items 19 and 20 which all students answered correctly, have relatively low average confidence rating (4.53) when compared to items 11 and 12 which were correctly answered by 26 and 27 students, respectively. The average confidence rating for these two relatively difficult items is 4.91. In such a case where relatively low confidence rating accompany correct answers, it is reasonable to assume that the students have an inaccurate perception of their knowledge. Such phenomenon, i.e., when correct answers are relatively rated low, is characterized by yule et al (1985: 47) as "non-confident correct answering".

Apart from non-confident correct answering, the students' ratings also indicated "very confident wrong answering" to use Yule's term. 25 students answered correctly item 18 and the average confidence rating indicated for this item was 4.84. Surprisingly, the average confidence rating indicated by the four students who got the item wrong was 5. Items 13, 16 and 22 are the other extreme examples of this same instance whereby relatively high confidence rating accompany wrong answers.

In such cases where wrong answers are relatively rated high, one might conclude that the students' perception of their knowledge is very misleading, because the students are tending to indicate thy are very 'sure', they are correct, in producing the answer when in fact, they are wrong.

Surely, in such instances where wrong answers are given high level of confidence it is a clear sign of fossilized inter-language. In other words, a high degree of fossilization leads to high level confidence in one's wrong answers. It is important to note that fossilizable structures tend to remain as potential performance, re-emerging in the productive performance of IL even when seemingly eradicated" (Selinker in Richards, 1974: 36).

According to Krashen, learning a language consciously as opposed to language acquisition (unconscious), the learner must know the rules of the language which serves as an editor or monitor of his linguistic output. Krashen has argued that the monitor acts as an editor particularly in reading and writing in second language because under these conditions there is time to go over and check the linguistic output. In his view, some language learners over use the monitor and become inhibited, others are overconfident and underuse it (in Stern, 1983: 403-4).

Asking the students to grade their 'confidence' is calling for maximum use of the monitor. When we look at Appendix C, we see that 13 students (44.8%) rated themselves "4" and above for their wrong answers. Conversely, 7 students (24.1%) graded their confidence between "3" and below. The remaining students (31%) indicated their confidence between "3" and "4". Perhaps, those students (that indicated "4" and above) who are overconfident and underused the monitor are poor monitor users in Krashen's terms. It seems reasonable to argue that a high level of confidence for wrong answers is a clear indication of fossilized errors, and hence overconfident students produce fossilized errors possibly because they underused the monitor. In other words, fossilized errors are produced when output is not monitored. Accordingly, *10, *15, *20 and *24 are instance of this situation (cf. Appendix B).

Meanwhile, in order to achieve native speaker competence, successful language learners cannot possibly have been taught this competence, since linguists are daily discovering new and fundamental facts about particular languages. Chomsky (1969: 68) expressing this similar view wrote:

> ...it must be recognized that one does not learn the grammatical structure of a language through 'explanation and instruction'beyond the most elementary rudiments, for the simple reason that no one has enough explicit knowledge about this structure to provide explanation and instruction.

On the other hand, according to Krashen's Monitor Hypothesis, learning, conscious knowledge, serves only as an editor, or a monitor of the learner's linguistic output.

Put in another way, for editing their output second language students use the learned, conscious knowledge, rather than the acquired.

Since this test is looking at the learned rather than the acquired system, students who indicated a high level confidence for their wrong answers could possibly be said to have a weaker learned system, but a better acquired system. In such cases where wrong answers are given with high level of confidence, surely, it indicates not only that the students are poor monitor users but also that they are having fossilized system. This is not only futher evidence of what yule et al called "non-confident correct answering" but also "very confident wrong answering" (Ibid., 477).

From Table 2, we could also observe that the students have unequivocally indicated that they are not at all sure of the answer they gave. Item 12 is a case in point.

Discussion

Paul L. Dressel (1976: 210), writing on the purposes of grade, noted that grades can be useful feedback to the student. Harald B. Allan (1965: 372), on the other hand, noted that "Nothing is more frustrating in the area of testing than to be given a test score, even in precentage or percentile terms, without a ready means of interpreting this score in terms of some immediately practical consequence." It is, however, cutomary while interpreting test results to treat two students that have identical scores as having performed similarly. Let us compare a few of the students who obtained the same overall test scores.

Table 3. Grade and confidence ratings of two students.

Students	% Correct	Confidence Ratings for Correct Ans.	Confidence Ratings for Wrong Ans.
S21	70	4.47	2.63
S29	70	4.42	4.25

According to the above table, both S21 and S29 have shown on the test relatively high confidence in the correct answers they gave. But their profile would not remain the same if we were to compare their average confidence ratings for wrong answers. It appears that S29 is an overconfident learner who is not at all aware of the fact that he does make mistakes. For sure, he has a fossilized system: for otherwise he wouldn't give such high level confidence for his wrong answers. Despite his high degree fossilization, this student could probably be communicatively efficient for his needs as opposed to S21. This is so because he is not inhibited by the fact that he makes mistakes.

In contrast, S21 is conscious of the answers: he knows when he is correct or not. He exhibits much less average confidence for wrong answers (2.63) in contrast to S29's 4.25. He, therefore, seems to be still learning, 'testing hypothesis,' 'making generalization,' etc., about the language he is learning. In addition, S21 could possibly be modest and introspective who always puts lower rating as opposed to the overconfident S29.

If, however, the only information available about these two students was their test scores, both getting 70, then they would tend to be considered as having performed at a similar level. Let us compare pairs of students who have obtained one of the highest scores, and the lowest ones.

Table 4. Grade and confidence Ratings of 4 students.

Students	% Correct	Average Confidence Ratings for Correct Ans.	Wrong Ans.
S1	85	4.83	3.75
S11	85	4.22	2.75
S13	44	4.50	4.00
S18	48	2.87	3.07

The 1st pair, S1 and S11 have the same percentage score; yet, S11 knows when he makes mistakes.

In the case of the 2nd pair, S13 S18, though they obtained almost identical scores, they show diametrically opposite tendencies in terms of their confidence ratings. The likely, if not altogether speculative, interpretation of this divergence might be that the 48% ercent correct score of S18 .

is a more genuine reflection of ability (in producing aspects of the language as they are presented in the test) than the 44 percent score of S13. In addition, S18 appears modest and introspective who always puts lower rating as opposed to the overconfident S13 who does not know when he makes mistakes.

From this, though more speculative, it maight be possible to generalize that students who were able to produce correctly, but had little confidence in their producation, may either hesitate, or avoid, or evade responding to a message so long as the communication situation does not demand them to produce. But they are still learning, that is, testing hypothesis, making generalizations, etc., about the language they are learning. It is also presumed here that such students manifest much less fossilized errors in their IL system in contrast to their peers.

Alternately, students who indicated high level of confidence in their ability, but showed low performance. (as in the case of S13) are unaware that their production is incorrect. Unlike the former ones, this sort of students could probably be communicatively effective with high degree of fossilization in thier inter-language system. In both cases, however, the students will undoubtedly come across some communication difficulties.

Conclusion

From the performance of the students and their confidence ratings it could be deduced that there is not a one-to-one correspondence between accuracy and confidence in answering the test items. In other words, the students neither rated "5" each of the items they got correct nor "1" all those items they wer wrong in.

It is a common experience that teachers usually review those test items which most students gave wrong answers to. However, when the students produce wrong answers with complete "false" confidence in a test item, then it does not mean that they are simply guessing the answers to the questions. If it were so, their confidence ratings would not be high. rather means that they are overconfident and are not cognizant that they are having trouble with that grammar item tested. Needless to say, it is a sure indication of the students' wrong perception of their knowledge on that particular test item. One reason for the cause of this "false" confidence could perhaps be high degree of fossilization. "It is not clear why many learners become arrested at certain inter-language levels" (Stenn, 1983: 410).

Apart from the possibility of being poor monitor users, such students might be by nature overconfident, less critical, and less introspective but probably communicatively efficient in contrast to their peers who are modest, and are still learning.

Even though exams force students to answer and therefore take risks, it seems improper to draw conclusions about 'risk-taking' from this experiment. Yet, we would like to risk one: A student who faces multiple choice item about which he is unsure (where incorrect guessing doesn't carry a penalty) is inclined to take risk and make appropriate judgement concerning the use of partial information when his motivational nature does not permit him to kep sufficiently calm in the exam.

We pay off on right answers and discourage wrong ones at every level of the teaching professon. It is certainly a good thing to be right but if we are so preoccupied with being right, we have no room for the students to make mistakes; we may rob them of their most important learning experience. People learn from their mistakes. An overemphasis on the importance of being right and insistence upon perfection may boomerang to discourage people from trying at all.

These students who indicated low level of confidence for their wrong answers are in the process of learning. Here the teacher should encourage the students to express their doubts on those items they were not certain, but recorded correct answers. Learners need a great deal more fredom to try, to experiment, to explore, to talk about. This was beautifuly stated by an American 5th grade student who wrote to his teacher after they had discussion about love in his classroom:

> "I was surprised when we talked in our class about love yesterday. I learned a lot of things and I found out about how lot of others feel. But I was surely surprised because I never knew you could talk about things in school that you didn't get grades for (Strom, 1971: 162).

This is to say, teachers should review not only the wrong answers, but also those items which the students are not sure about. A good teacher is able to influence both student feeling and achievement in positive ways.

Finally, the forgoing observations have evident educational implication, and hence more systematic research specifically focussed on the issues would be highly worthwhile. Particularily, the application of the 'Theory of Signal Detection (TSD), effectively used by Yule et al (1985) would help us more in understanding of how individuals cope with uncertainity in communication and in language learning process.

References

Allan, Harold B. **Teaching English as a second language: A book of Readings**. New Delhi. McGraw-Hill, 1972.

Carroll. J.B. 'The prediction of success in intensive foreign language training.' **In Training Research and Education**, ed. R. Clazer Pittsburgh: University of Pittsbursh press, 1962.

Chomsky, Noam. **Current Issues in Linguistic Theory**. Hague: Mouton & Co., 1969.

Dressel paul L. **Handbook of Academic Evaluation**. London: Jossey Bass publishers, 1976.

Ebel, Robert L. **Essentials of Educational Measurement**. New Jersey: Prentice-Hall, Inc., 1972.

Ingenkamp, Karllreinz (ed) **Developments in Educational Testing** (vol. 1). New York Gordon and breach Science Publishers, 1966.

Johnson, Keith (ed) **Communication in the classroom applications and methods for a communicative approach** Hong Kong: Longman, 1981.

Lambert W.E, and R.C Gradner. **Attitudes and Motivation in second Language learning**. Rowley, Mass Newbury Huse, 1972.

Oller, J.W., A. Hudson, and P.F Liu "Attitudes and attained proficiency in ESL: a socialinguistic study of native speakers of chinese in the united states". **Language Learning**. Special Issue No. 4, pp. 1-27.

Pimsleur, P. **Language Aptitude Battery**. New york. Harcourt, Brace and world, 1966.

Richards, Jack C, (ed) **Error Analysis: Inspections on second language acquisition**. London: Longman Group Ltd. 1974.

Stern, H.H. **Fundamental concepts of language teaching**. Great Britain: Oxford university Press. 1983.

Strom, Robert D. **Teachers and the Learning Process**, New Jersey: Prentice-Hall, Inc., 1971.

Yule George, Yanz Jerry L, and Tsuda, Atskuo. "Investigating aspects of the language learner's confidence: An application of the theory of signal detection." **Language learning**. Vol. 35, No. 3, 1985.

Appendix A Confidence ratings of each student for ever item tested

Students	Phrasvos					Prepositions										Quantifiers					Tense				Reported spech		
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27
1	3x	4	3x	4	4	5	5	5x	5	5	5	5	5	5	5	5	5	5	5	5	5	4	4x	5	5	5	5
2	3x	3x	4x	4	4	4	5	4	4x	5	5	5	4x	4x	5	5	5	5x	5	5	4x	4x	4x	4x	4	5	5
3	3	3x	5	5x	5	5	3	5	5	5	3x	5	5x	2x	3x	5	5	5x	5	3	3	5x	5x	3	5	5	5x
4	5x	4x	5	4	4	5	5	5	5	5	4	5	5	5	4	5	3	5	5	5	4x	4x	4x	4x	4	4x	4x
5	5x	5x	5	5	1x	5	5	5	5	5	5	5	3x	5x	5	5	5x	5	5	5	5	5x	5	3x	5x	3x	3x
7	3x	3	5	3x	4x	5	5	4x	3	4	5	5	5x	4	4x	4	5	5	5	5	3	1x	4x	4x	5	5	4
8	2x	1x	1x	4x	4x	4x	5	1x	3x	2x	4x	5	5x	2x	2x	5	5	5	5	3	3x	2x	5x	3x	4x	5x	2x
9	3	4x	4	4x	5	5	4	5	5	4	5	5	4x	4x	4	5	5	5	5	5	4	4x	4	4	5	5	5x
10	5	5	5	5	5	5	5	5x	5	5	5	5	5x	5	4	5	5	5	5	5	5x	5	4	4	5	5x	5x
11	4	4	5	5	4	5	4	3x	2	4	5	5	3	4	5	3	2	5	5	5	5	2x	3x	3x	5	5	5
12	4x	4	5	3x	4	5x	4	5	5	4	5	5	4	5x	4	5	3	5	5	5	4x	5x	3x	5x	5	5	5
13	4x	4x	2x	3x	3x	4	5x	5	5	5	5	5	5x	5x	5x	4	4	5x	5	5	4	3x	4x	4x	3	4x	3x
14	5x	5	4	5	5x	5	5	3	4x	5	5	5	3	5	5	5	3	4	5	5	5	5	4x	4x	5	5	5
15	5x	4x	4	5x	5	5	5	5	5	4	5	5	5x	4x	4	5x	5	5	5	5	4	4x	3	5x	5x	5	4
16	3x	1x	2	1	1	4	2x	3x	5	4	5	4	5x	5	4	5	3	5	5	4	1	1x	1x	1x	5	5	5
17	5x	4	4	2	1	4	4	1	4x	4	5	1x	4x	3	2	5	5	5	4	5	4	4x	3x	3x	5x	5	4x
18	3x	3x	4	5x	2x	3	3x	1	2x	4x	4x	1x	3	3	2	5	5	5	5	4	2	3x	3x	3x	1	4x	3x
19	5x	3	1x	5	5	5	5	5	5	5	5	5	5x	5	5	5	3	5	5	3	5	5x	5x	5x	5	3x	5x
20	5	5	5	5	5	5x	4	5x	5	5	5	3x	5	3x	5x	5	5	4	5	5	5	3	5x	5x	5	5	4
21	3x	4	4x	5	5x	5	4	5	5	5	4	5	3x	5	5	5	5	4	4	4	1	1x	3x	1x	5	5	1x
22	4x	3x	4	4	4	4	4	5	4	3	5	5	5x	4	5	5	4	5	3	5	3	4x	4x	4x	5	4x	1x
23	5	5	4	5	3	5x	5	4	5	5	5	5	5x	4x	5x	5	5	5x	5	5	5	5	4	5x	5	5x	3x
24	5	5	5	5	5	4	5	5	5x	5	5	5	5x	5	5	5	5	5	5	5	4	5x	5	5x	5	5	5x
25	1x	1x	2x	1x	1x	5	5	5x	4	5	4	5	5x	2x	5	5	3x	5	5	5	5	5	3	4	4x	5x	3x
26	4x	3x	3x	2x	2x	4	2x	3x	4x	3	4	5	5x	1x	3x	4	3	4	3	3	3x	1x	1x	3	4x	2x	3x
27	3x	4x	3	2x	3	1x	3x	1x	2x	2x	5	5	5x	1x	3x	5	1x	5	5	4	5	2x	5x	1x	5	5x	4x
28	4	5	4	5	4	4x	5x	5	5	5x	5	5	5x	2x	4x	5	4	5	4	4	4	4	3	3	5	4	4x
29	4x	4	5	1x	2	4	5	5	5x	5	5	5	5x	5	5	5	5x	5	3	5	5	4x	5x	5	5	5x	5

Appendix B

Score and mean confidence for actually right, and wrong answers.

Student	Score	No. of Correct	No. of Wrong	Rectaings for Correct	Wrong
1	85	23	44	4.83	3.75
2	59	16	11	4.63	3.91
3	63	17	10	4.41	4.10
4	70	19	8	4.63	4.13
5	59	16	11	5.00	3.91
6	74	20	7	4.75	3.71
7	67	18	9	4.44	3.56
8	26	7	20	4.71	2.80
9	78	21	6	4.57	4.17
10	81	22	5	4.86	5.00
11	85	23	4	4.22	2.75
12	70	19	8	4.58	4.27
13	44	12	15	4.50	3.93
14	81	22	5	4.64	4.40
15	67	18	9	4.61	4.67
16	70	19	8	3.84	2.13
17	67	18	9	3.67	3.67
18	48	13	14	3.31	3.07
19	70	19	8	4.68	4.25
20	74	20	7	4.70	4.71
21	70	19	8	4.47	2.63
22	70	19	8	4.26	3.63
23	70	19	8	4.74	4.63
24	81	22	5	4.91	5.00
25	56	15	12	4.67	2.75
26	41	11	16	3.64	2.63
27	37	10	17	4.50	2.65
28	74	20	7	4.35	4.14
29	70	19	8	4.63	4.25