9-2002

An Assessment of Basic Computer Proficiency Among Active Internet Users: Test Construction, Calibration, Antecedents and Consequences

Eric T. Bradlow
University of Pennsylvania

Stephen J. Hoch
University of Pennsylvania

J. W. Hutchinson
University of Pennsylvania

Recommended Citation

An Assessment of Basic Computer Proficiency Among Active Internet Users: Test Construction, Calibration, Antecedents and Consequences

Abstract

The purpose of this article is to describe our efforts to create a test of basic computer proficiency, examine its properties using parametric test scoring methods, and identify some antecedents and consequences that accompany differences in performance. We also consider how much insight people have into their level of knowledge by examining the relationship between our tested measure of computer knowledge and self-rated knowledge scores collected at the same time. This research also adds to the large body of existing empirical work on computer literacy in the student population, by looking at computer literacy in a more general sample of the Internet-using population. A further purpose of this research, as a result, is to make our dataset available for future research.

Keywords
calibration, computer proficiency, parametric test scoring

Disciplines
Business | Educational Assessment, Evaluation, and Research | Marketing | Science and Technology Studies | Statistics and Probability

An Assessment of Basic Computer Proficiency Among Active Internet Users:

Test Construction, Calibration, Antecedents, and Consequences

By

Eric T. Bradlow

Stephen J. Hoch

J. Wesley Hutchinson

Eric T. Bradlow is Associate Professor of Marketing and Statistics, Stephen J. Hoch is John J. Pomerantz Professor and Marketing Department Chair, and J. Wesley Hutchinson is Stephen J. Heyman Professor at the Wharton School of the University of Pennsylvania. All correspondence on this manuscript should be addressed to Eric T. Bradlow, ebradlow@wharton.upenn.edu, 3620 Locust Walk, Philadelphia, PA 19104-6371.

Abstract

The purpose of this article is to describe our efforts to create a test of basic computer proficiency, examine its properties using parametric test scoring methods, and identify some antecedents and consequences that accompany differences in performance. We also consider how much insight people have into their level of knowledge by examining the relationship between our tested measure of computer knowledge and self-rated knowledge scores collected at the same time. This research also adds to the large body of existing empirical work on computer literacy in the student population, by looking at computer literacy in a sample, albeit self-selected, of the internet using population.

A summary of our empirical findings, based on a sample of 1520 respondents' answers collected from the Wharton Virtual Test Market suggest: (1) the test instrument was approximately unidimensional, (2) people are moderately calibrated with respect to their basic computer proficiency relative to that of others, but are not well-calibrated regarding their knowledge of different subdomains of computer proficiency, and (3) that various antecedents of computer knowledge (e.g. on-line experience, familiarity with technology) were significant predictors of objective knowledge, self-report knowledge, and calibration, while in addition these measures were all significantly related to consequences such as number of online purchases, and concern over buying from an internet retailer.

1. Introduction

Computers are a ubiquitous part of everyday life, though 20 years ago most people had little experience with any kind of computer, mainframe or personal. Most people over 50 years old had no exposure to computers in secondary school or college. Today, children begin using computers for play and work, in schools and at home, at a very early age. E-mail is a very common form of communication for young and old, but until the advent of the first internet browsers in the mid 1990's, few people had access to it. In recent years, advances in software tools and computing power has made the entire experience of using a computer much more user friendly; the minimum knowledge threshold to effectively use a computer is much lower today than it was even a few years ago. In fact, such advances have been the result of a concerted effort on the part of computer manufacturers, on-line service providers, and on-line retailers to increase the installed base of potential consumers.

In conjunction with the increased importance that computers play in everyday life, the societal norms for computer literacy have changed. As an example, many states have instituted computer proficiency examinations as a requirement for graduation, to ensure that all high school graduates obtain basic computer mastery. For example, beginning with the high school class of 2000, the state of North Carolina instituted a mandatory computer proficiency exam (http://www.ncpublicshools.org/testing/). As we concur that understanding basic computer mastery has potential importance, we address it in this research. Therefore, we set about to construct a test that would provide a measure of basic computer ability and knowledge in the general population of computer users,

without being constrained to evaluate any specific educational program. A second goal was to demonstrate how parametric psychometric models could inform test developers about the quality of test construction. A third goal was to provide a test instrument and empirical results for a rapidly changing aspect of society that might aid future research directed at tracking these changes.

Then, after simultaneously collecting self-reported knowledge and internet behaviors, we link our test scores to people's perceptions of their own mastery level, collected demographics, and internet outcomes. Such an analysis extends the contribution of this research beyond one of pure measurement instrument construction and assessment, to one that allows us to hypothesize antecedents and consequences (e.g. internet behaviors) of obtaining basic computer proficiency. In addition, as much empirical research has focused on basic computer mastery at the student level, or at obtaining advanced mastery skills (McIrney, McIrney, and Marsh, 1997), we focus here on basic mastery for a more broadly based internet using population.

The remainder of this article is laid out as follows. In Section 2, we describe the method used to construct our test of basic computer proficiency. In addition, we describe the population to whom our test was administered as well as the other variables collected along with our objective test items and self-reported knowledge scores. Section 3 contains the results of our empirical analyses including: tests for unidimensionality of the instrument (using factor analysis and a testlet item response theory (IRT) model), confirmation that the battery of items is testing for *basic* computer proficiency (using the estimated IRT difficulty parameters), an assessment of people's knowledge calibration by relating their observed and self-report test scores, and finally an assessment of the

relationship between these scores and antecedent and consequence variables.   We conclude in Section 4 with a discussion.

2. **Method**

2A. **Test Construction**

Although there is little extant research examining the substantive and statistical characteristics of basic computer proficiency tests and test items on non-student populations, there are a large number of objective tests and self-report checklists that are used by high schools and universities. We reviewed several such instruments accessible on the world-wide-web (WWW) to help generate basic computer proficiency test items that we would use for our population of interest. We were also guided by the standards set for high school students by the National Educational Technology Standards Project (International Society for Technology in Education 1998) and by our own experience regarding computer-related competencies in universities, business practice, and consumer behavior. Because of our focus on basic proficiency, items were sought that reflected common knowledge or facts useful in a wide variety of common situations. We specifically avoided items that were directed at advanced levels of experience and proficiency.

We developed a pool of items in nine subdomains of computer proficiency: terminology, file management, word processing, spreadsheets, data bases, printing, email, the internet, and information search. A small convenience sample was used to pretest and narrow the pool. The final test consisted of three items in each of the nine subdomains for a total of 27 items (see Table 1). Four of our subdomains (terminology, word processing, spreadsheets, and data bases) are directly included in most state tests of basic

student computer mastery, while three of them (printing, email and the internet) are included in part under the term telecomputing. We added file management and information search as they are relevant to our population of interest, which includes computer users involved in business-related activities.

2B. **Sampling**

Data to form the basis of our research were collected from Internet users who participated in an on-going panel called the Wharton Virtual Test Market (WVTM) . Each year from 1997-2001, approximately 12,000 respondents filled out a base WVTM survey as well as up to six additional on-line surveys per year. Each survey contains questions related to different aspects of online and offline behavior and attitudes. The surveys take approximately 15-20 minutes to complete. Panel members were recruited through two methods: banner advertisements and opt-in e-mail lists. There were two financial inducements: (1) a 25% chance to win a free $10 phone card; and (2) a much lower unspecified chance to win a $250 and $500 lottery.

Each year panelists were recruited sequentially over about a 6-month period. At various points in time over the 6-month period, the demographic characteristics of our sample were compared to the US Census' estimates of the composition of the online population. When there were significant deviations between the two samples, the WVTM altered both where recruitment banners were run, and which e-mail lists were purchased in order to over-sample under-represented sub-populations. The full sample did not significantly differ from the US Census estimates (all χ^2's had p>.2). For the present research, we used data from panelists who participated in the WVTM during the first 6 months of 2001. Because of the specific character of the recruitment devices used

in online survey research, it is not possible to compute a meaningful response rate. We do know, however, that of the people who started to complete the survey, about 85% finished.   Although this survey does not represent the entire US population as a whole, or even the entire online population, it is representative of a significant population of the internet users.   Recent figures have suggested that 50% of the US population owns computers, with over 90% of these persons being on-line. Figures have also shown the fraction of on-line users to computer owners as increasing, with the possibility of one day being above 1 (due to sources of internet access other than personally owned computers). Thus, we believe this a relevant population for study.

The computer proficiency test was completed by a sub-sample of the panelists during one of the normal data collection periods. An e-mail invitation was sent to 2,000 panelists to complete the questionnaire and 1,520 completed it for a 76% response rate. The demographics of this sub-sample also closely matched the US Census estimates of the characteristics of the online population.

2C. Online and Offline Behavior and Attitudes Survey

As part of the WVTM, panelists completed an online survey. Part of the survey asked about online purchase and search behavior. Respondents were asked how many online purchases they had made in the immediately preceding 6 months and then asked to describe their last purchase and indicate the amount paid. They also indicated whether they had searched and/or purchased a product in the following categories: leisure travel, automobiles, financial services, computer systems, software, books, clothing, groceries, consumer electronics, music, videos, toys, and housewares.   They also provided information about offline purchasing activities including buying from printed catalogs

and retail stores. Panelists answered a series of attitude questions about technology products including email, the desire for new electronic gadgets, and their own technology savvy. They also expressed their attitudes about whether they would only order products from well-known companies and their concerns about privacy and transaction security. Panelists indicated the speed of the Internet connection and other activities for which they used the Internet often including: work/research, news/current events, entertainment, banking, health care, travel, and government. Finally, panelists provided information on a comprehensive set of demographic characteristics including: age, education, occupation, income, race, gender, marital status, and number of children.

3. **Results**

The results are divided into four parts. First, we report the statistical characteristics of the items in the test. These results demonstrate that the items measure a single construct, that discrimination of most of the items is focused at a basic level, and that smaller, more efficient subsets of test items can deliver nearly the same amount of information. Second, we examine the relationship between tested and self-rated computer proficiency. These analyses can reveal how much insight people have into their own computer proficiency and knowledge. Third, we present a set of regression analyses where we examine the relationship between tested proficiency, self-rated proficiency, and the difference between tested and perceived computer proficiency (knowledge calibration) and a set of antecedent variables including prior online experience and personal demographic characteristics. Finally, we consider the influence of computer proficiency on online behaviors and attitudes (the consequences).

3A. Item Characteristics

The results from several different analyses demonstrate that the 27 items measure a single construct. First, raw scores (i.e., number correct) on the nine subtests were factor analyzed. Only two factors had eigenvalues greater than 1 and a chi-square test of the hypothesis that more than 2 factors are required to explain the correlations among the nine subscale scores failed (χ^2 = 17.3, df = 19, p =.57). The first factor accounted for approximately twice as much variance as the second and had the highest loading for 7 of the 9 subscale score variables. The two subtests that loaded highest on the second factor were Email and Information Search, suggesting the possibility that internet-related knowledge was distinct from more basic knowledge about computer usage. However, the Internet subtest loaded highest on factor 1 and 2 of the 3 Terminology questions were internet-related. Therefore, we conclude that the data were "approximately" unidimensional. Further evidence of unidimensionality is discussed subsequently in the context of the results of estimating item response theory (IRT) models from these data.

To assess the statistical properties of our test, we *initially* fit the standard three-parameter item response theory model (3PL IRT, Birnbaum 1968) to the WVTM data given by

$$p(y_{ij}=1) = c_j+(1-c_j)*\text{logit}^{-1}[a_j(\theta_i-b_j)]$$

where y_{ij} is the score for item j received by examinee i, $p(y_{ij} = 1)$ is the probability that examinee i answered item j correctly, θ_i is the (latent) ability of examinee i, a_j denotes the discrimination (slope) of item j, b_j is the difficulty of item j, and c_j is the pseudo-chance

level parameter of item j. For a full explication of this model see Hambleton & Swaminathan (1985), Lord (1980) and Wainer & Mislevy (1990). We note that this approach to test property exploration, using a parametric model, is by no means exhaustive, and is supplemented in part (below), by more traditional measures. Nor in fact is it new to the literature as an application of such procedures, as applied to the North Carolina Test of Computer Skills, appears in Wang, Bradlow, and Wainer (2002).

The parameters estimated from our computer proficiency data are given in Table 1. In general, item discrimination is high. Eighteen items have values of a greater than one and only one value is close to zero. Item difficulty is moderate to low. Sixteen items have negative values of b. This is consistent with our goal of focusing the discriminating power of the test on basic computer proficiency. Finally, there is not a high level guessing. Sixteen items have values of c that are less than $1/k$ (where k is the number of response alternatives for the item). Thus, from the perspective of having items with low to modest guessing, high discrimination, and being "located" at the basic proficiency part of the scale, we felt that our test items met those criterion.

One basic concern about applying the standard 3-PL model to our data lies in the assumption of conditional independence among the items. This is especially true in this case as we designed the test to contain 9 "testlets" (Wainer & Kiely, 1987), in this case areas of subdomain expertise for computer proficiency. To assess the reasonableness of this assumption, we fit the so-called Testlet IRT model initially proposed by Bradlow, Wainer, and Wang (1999). The Testlet IRT model extends the basic Birnbaum model to allow for increased dependence of items within the same subscale (in our case computer proficiency domain), by modeling

$$p(y_{ij}=1) = c_j + (1-c_j) * \text{logit}^{-1}[a_j(\theta_i - b_j - \gamma_{id(j)})]$$

where $\gamma_{id(j)}$ is an effect for person i on the testlet containing item j, d(j). As in Wang, Bradlow, and Wainer (2002), this in essence provides subscale estimates of ability, $\theta_i - \gamma_{id(j)}$, for person i on subscale d(j). The model results revealed that there was little to moderate excess dependence between items in the same subtest. The variance of the testlet effects ranged from a low of 0.104 for Information Search to a high of 0.697 for Email. When compared to the variation in the underlying abilities (fixed to N(0,1) to identify the model), we see that the largest testlet effect variance is of modest level, indicating that the primary source of variation is the underlying abilities. This supports the conclusion that "approximately" a single ability is tested (as in the factor analysis results), or at least that the degree of non-unidimensionality is not highly concentrated on any single sub-domain.

Overall, the test appears to be reasonably constructed by traditional criteria (e.g., Holland & Wainer, 1993; Stocking and Swanson, 1998). For future research, however, it appears that shorter, more efficient versions of the test are possible. To examine this issue the observed Fisher information was computed for each item as follows.

$$\hat{I}_j = \Sigma_{i=1}^{I} \{\hat{a}_j^2 * [\exp(\hat{t}_{ij})/(1+\exp(\hat{t}_{ij}))^2 - y_{ij}\hat{c}_j \exp(\hat{t}_{ij})/(\hat{c}_j + \exp(\hat{t}_{ij}))^2]\}/I,$$

where $\hat{a}_j, \hat{b}_j, \hat{c}_j$ are the estimated item parameters for item j, $\hat{\theta}_i$ is the estimated ability

for respondent i, and $\hat{t}_{ij} = \hat{a}_j(\hat{\theta}_i - \hat{b}_j)$ is the estimated latent linear score for person i and

item j. The results reveal that 75% of the information in the test is captured by the top 10

items. We further found that 69% of the information is captured by the 9-item set formed

by the top item in each subtest, and the 18-item set formed by the top two items in each

subtest captures 94% of test information. Finally, all four True/False items were among

the bottom 10 in information (and had either discrimination or guessing problems as

well). This confirms the common testing practice of avoiding True/False items is a

useful guideline in this domain as well. In summary, for the purposes of scale construct

reliability in this initial foray, having "excess" items may have been of some value, but

more efficient tests based on the results reported in Table 1 should be sufficient for most

purposes in future research.

3B. Knowledge Calibration

The relationship between objectively measured knowledge and subjective beliefs

about one's own knowledge has been studied in a variety of ways and is typically called

knowledge calibration (e.g., Alba and Hutchinson, 2000). We computed the relation

between tested (i.e., measured by our 27 item battery) and self-reported computer

proficiency in two different ways: within and across respondents. For the within person

measure of calibration, for each respondent we first calculated their percentage correct

for each of the 9 knowledge subscales. We then computed a within person correlation

between these subscale percent correct scores and the self-rated knowledge ratings

(measured on a 0-3 Likert scale) for the same subscales. This analysis indicated only a

modest association between tested and self-rated computer proficiency, an average

correlation across respondents of 0.195 (median = 0.222). The average and self-rated proficiency scores, shown in Table 2 columns 1 and 2 provides another view of why the within person correlations are not substantial. Respondents were quite overconfident in the knowledge of the internet, whereas they were underconfident in their knowledge of computer terminology and file management, at least for the battery of items that we constructed.

We also examined calibration across respondents. For each respondent we computed their measured proficiency across all 27 items and the average self-rated proficiency for the 9 knowledge domains. The across-person correlation between tested and self-rated proficiency was 0.51 suggesting that persons were better calibrated in what they knew in comparison to others, as opposed to what they knew on an objective level. In addition we computed across person correlations between tested and self-rated proficiency at the subscale level. Those correlations appear in Table 2 column 3. Calibration is highest for spreadsheets and lowest for information search and e-mail; most importantly they do vary widely suggesting differential ability of persons to assess their relative knowledge across subdomains. A subsequent analysis, in which we correlated an "error-free" ability subdomain estimate of θ_i-$\gamma_{id(j)}$, with the self-report scores, indicated an identical ranking pattern as Table 2, column 3 with correlations roughly 10% higher.

3C. Antecedents of Computer Proficiency

In this section, we examine differences in tested and self-rated computer proficiency based on level and type of online computer experience, demographic characteristics, and self-rated technology expertise. In labeling these variables "antecedents" we do not mean to make any strong claims about causality. The measures

were temporally prior to our current measures and logic or common sense suggests they might exert some causal effect on computer proficiency. In some cases, it is possible that the relationship could be bidirectional or mediated by some third (possibly unobserved) variable. Thus, we consider these to be exploratory analyses. We used four dependent variables: (1) tested proficiency, (2) self-rated proficiency, (3) a measure of knowledge calibration, and (4) the estimated computer ability parameter, $\hat{\theta}$, from the IRT testlet model analysis.

Table 3 contains the results. The significant effects (at the 0.05 level) are shaded in gray. As can be seen, we see a very similar pattern of results for tested and estimated proficiency. Respondents who have more online experience tend to demonstrate greater computer proficiency. This is not particularly surprising. Each year of additional online experience results in about ¾ of a point improvement in tested performance, which is about 17.6% of a standard deviation (4.25 points). Moreover, people who spend more hours online each week also score higher, about .16 points for each additional hour each week. Alternatively, level of email usage is not a significant predictor of objective computer proficiency, whereas it is for self-rated knowledge. Respondents also rated their level of technology familiarity and desire for the latest new electronic products. People who considered themselves more technology savvy did indeed display greater computer proficiency and a higher self-assessment of their knowledge. However, those who like to have the latest electronic gadgets actually scored worse on the computer proficiency test, yet rated their own self-knowledge as higher.

There are also some significant differences in computer proficiency amongst demographic groups. In terms of occupation, students and people working in technical

fields scored better. There was a small but statistically significant positive effect for income. African Americans scored over 1 ¼ points lower, though the standard error is quite large. Not surprisingly, there was a positive relation between computer proficiency and education. Younger people scored higher than older people as well. These results parallel the widely reported relationships between the same demographic variables and access to computers and the internet (often called the "Digital Divide"; National Telecommunications and Information Administration, 1999). They provide some empirical evidence that the differences extend beyond access to proficiency.

An interesting comparison emerges when we examine the pattern of coefficients for the tested versus self-rated proficiency dependent variables. Generally the coefficients are of the same sign. The two notable exceptions are for African American and desire for new electronic gadgets, where in both cases there is a negative coefficient for tested proficiency and a positive one for self-rate knowledge.

In the final analysis, we examined what factors influence knowledge calibration. To do this, we first needed to compute the difference in tested and self-rated computer knowledge. We computed standard variates (z-scores) for both tested and self-rated computer knowledge and then the squared difference between the two. The larger is this calibration quantity, the less closely aligned is a respondent's self-rated computer knowledge with their tested knowledge. We also examined the raw difference between the two z-scores, which gives an indication of whether the error is one of over- or under-confidence in tested knowledge. We found that people with more online experience actually were less accurate in their assessment of their computer knowledge. Surprisingly, this occurred because more online experience led people to under-estimate

their tested computer knowledge --- people with more online experience display greater computer proficiency, but for some reason their self-ratings did not reflect that greater proficiency. The opposite occurred for the technology familiarity variable. People who rate themselves as having more familiarity with technology display greater computer proficiency, but not as much as they thought. Students were less well calibrated while professionals showed better calibration. Higher income groups also showed better calibration whereas better educated groups showed worse calibration. These findings certainly suggest that knowledge and assessment of one's knowledge in the computer proficiency domain is an interesting area for future research.

3D. **Consequences of Computer Proficiency**

We also were interested in whether computer proficiency influenced online purchase behaviors and attitudes towards privacy and security. For this analysis we regressed the variables listed in Table 4 onto both self-rated and tested computer proficiency. The first 3 variables were measured on 5-point agree (5)---disagree (1) scales.

As indicated in Table 4, people who score higher on the computer proficiency test are generally more trusting of the Internet. They indicate that they are more willing to buy from companies that they do not know that well and they are less concerned about both privacy and monitoring of their online activities and the security of their online transactions. It is an interesting question as to which of these opinions are correct. Are less knowledgeable respondents displaying a bit of Internet paranoia or are more knowledge respondents being a bit too cavalier? It is true that Internet users generally report few big problems when it comes to online transactions; in our data a majority

indicate that they never have had a bad experience with an online transaction or purchase. At the same time, the potential for problems with privacy and security do exist and maybe only time will tell. At this point, however, consumers have no greater than a $50 liability for a stolen credit card number and few hassles, except in the very rare case of complete identity theft (Weber 2001). Table 4 also shows that people with greater computer proficiency, both tested and self-rated, are more likely to make online purchases. Computer proficiency is however unrelated to purchases initiated by clicking on a banner advertisement.

4. Discussion

Our empirical findings suggest the pool of test items used in this study yield an approximately unidimensional measure of basic computer proficiency. Furthermore, our sample of internet users people were moderately calibrated with respect to their basic computer proficiency relative to that of others, but were not well-calibrated regarding their knowledge of different subdomains of computer proficiency (albeit this result may be due in part to the idiosyncracies of the items we used). Finally, various antecedents of computer knowledge (e.g. on-line experience, familiarity with technology) were significant predictors of objective knowledge, self-reported knowledge, and calibration, and these latter measures were significantly related to consequences such as number of online purchases, and concern over buying from an internet retailer.

More generally, tracking basic computer proficiency as it develops worldwide is an important task for educational research in the coming years. It provides a rare opportunity to study the acquisition of essentially the same knowledge and skills across

the full demographic spectrum in a variety of learning situations. An overarching goal in

the research reported here has been to provide impetus to this task.

References

Alba, J. W. and J. W. Hutchinson (2000), "Knowledge Calibration: What Consumers Know and What They Think They Know," *Journal of Consumer Research*, 27 (September), 123-156.

Birnbaum, A. (1968), "Some Latent Trait Models and Their Use in Inferring an Examinee's Ability", In F.M. Lord and M.R. Novick, *Statistical Theories of Mental Test Scores* (chapters 17-20). Reading, MA: Addison-Wesley.

Bradlow, E.T., Wainer, H., and Wang, X. (1999), A Bayesian Random Effects Model for Testlets, *Psychometrika*, 64, 153-168.

Hambleton, R.K., and Swaminathan, H. (1985). *Item Response Theory: Principles and Applications.* Boston: Kluwer.

Holland, P.W. and Wainer H., (1993) *Differential Item Functioning*, Hillsdale, NJ, Lawrence Erlbaum Associates.

Lord, F.M. (1980), *Applications of Item Response Theory to Practical Testing Problems*, Hillsdale, NJ: Lawrence Erlbaum Associates.

McIrney, V, McIrney, D.M., and Marsh H.W. (1997). "Effects of Metacognitive Strategy Training Within a Cooperative Group Learning context on Computer Achievement and Anxiety: An Aptitude-Treatment Interaction Study," *Journal of Educational Pychology*, 89 (4), 686-695.

National Telecommunications and Information Administration (1999), "Falling Through the Net: Defining the Digital Divide." Technical Report.

Stocking, M.L. and Swanson, L (1998), "Optimal Design of Item Banks for Computerized Adaptive Tests", *Applied Psychological Measurement*, 22, 271-279.

Wainer, H. and Kiely, G. (1987), Item Clusters and Computerized Adaptive Testing: A case for testlets, *Journal of Educational Measurement*, 24, 185-202.

Wainer, H and Mislevy R.J. (1990), Item Response Theory Item Calibration and Proficiency Estimation. Chapter 4 in Wainer, H., Dorans, N.J., Flaugher, R., Green, B.F., Mislevy, R.J., Steinberg, L., and Thissen D. *Computerized Adaptive Testing: A Primer* (pps 65-102). Hillsdale, NJ, Lawrence Erlbaum Associates.

Wang, X., Bradlow, E.T., and Wainer, H. (2002), "A General Bayesian Model for Testlets: Theory and Applications", *Applied Psychological Measurement*, 31, to appear.

Weber, T.E. (2001), "What Do You Risk Using a Credit Card on the Net?," <u>Wall Street Journal</u>, B1 12/10/01.

Table 1
IRT Model Results and Item Descriptions

Test Item	Parameter (S.E.)			Info.
	a	*b*	*c*	
Terminology				
1 A(n) _____ is a small program that can provide animation and sound clips on a Web site. (Java bean; Gif; Code Warrior; Applet)	1.37 (.31)	-.08 (.32)	.34 (.10)	.168
2 Which of the following is not an operating system? (Unix; Windows NT; Mac OS 9.0; Netscape)	2.58 (.43)	-.44 (.10)	.10 (.05)	.691
3 My computer "desktop" is: (the work space within any program I have opened; the main screen of the operating system, which may contain wallpaper, program icons, task bars, and open windows; the space where my computer; monitor; keyboard; and mouse are loc	1.38 (.18)	-1.82 (.31)	.18 (.15)	.140
File Management				
4 For most computer operating systems, a "folder" is the same thing as: (a directory; a file; an application; a network)	2.71 (.41)	.44 (.06)	.07 (.02)	.677
5 In the Windows operating system, the program to search for specific files is called: (ScanDisk; Internet Explorer; Windows Explorer)	1.75 (.42)	-.42 (.35)	.45 (.12)	.192
6 After you move a file to the recycle/trash bin, you: (can never retrieve it again; can retrieve it with a special program; 2:can only retrieve it if you have not emptied the recycle/trash bin)	.59 (.32)	-.61 (1.57)	.75 (.12)	.012
Word Processing				
7 Most word processors can: (check spelling; check spelling and grammar; check neither)	1.32 (.49)	-.24 (.63)	.52 (.13)	.105
8 In the word processing program, Microsoft Word, a file named "File" is saved as: (File.word; File.com; File.txt; File.doc; File.wrd)	2.22 (.48)	-.38 (.17)	.12 (.08)	.525
9 Using a word processor, the sentence, "The sly fox jumped over the fence," was edited with the following actions: (True; False)	2.39 (.65)	.85 (.17)	.65 (.02)	.092
Spreadsheets				
10 When entered into a spreadsheet cell, "SUM(B2:B5)" is called a: (label; value; function; cell)	2.54 (.54)	.80 (.09)	.26 (.03)	.302
11 Which is an absolute address? (C4; C4; 10700 – 105 Ave.; None of the above)	1.45 (.25)	1.64 (.12)	.03 (.02)	.166
12 In most spreadsheets, the cell labeled E8 is in which row and which column? (5th row, 8th column; 8th row, 5th column; Don't know)	1.61 (.31)	.98 (.11)	.10 (.04)	.237
Databases				
13 Which of the following cannot be entered into an alphanumeric field in a database? (dog; dog99; 99; none of the above (all are alphanumeric))	1.21 (.25)	-.25 (.32)	.21 (.11)	.189
14 Databases can manage more information than spreadsheets because most information is stored on a disk drive rather than in RAM. (True; False)	.02 (.01)	-.77 (1.56)	.07 (.03)	.000
15 Which of the following is a data base program? (Excel; Word; Access; PowerPoint)	2.15 (.29)	.27 (.07)	.06 (.03)	.547

Table 1
IRT Model Results
(Continued)

Test Item	Parameter (S.E.)			Info.
	a	*b*	*c*	
Printing				
16 Which of the following printer characteristics are NOT controlled by the Page Setup window: (landscape vs. portrait orientation; maximum font size; page size; margins)	1.86 (.29)	-.10 (.11)	.07 (.05)	.467
17 Which of the following types of printers has the highest print quality (i.e., resolution)? (dot matrix; ink jet; laser jet; bubble jet)	1.02 (.21)	-2.00 (0.71)	.31 (.22)	.077
18 Of the following, which is most likely to cause printer errors? (insufficient computer memory; insufficient printer memory; insufficient computer speed; insufficient printer speed)	.67 (.11)	.15 (.28)	.09 (.07)	.084
Email				
19 The ability to exchange e-mail messages containing sound, picture, and video files is: (currently available; likely to be available soon; not technologically feasible; don't know)	1.04 (.29)	-2.23 (.80)	.33 (.23)	.067
20 E-mail messages are private and can only be read by the sending user and the receiving user. (True;False)	1.06 (.44)	-.09 (.90)	.51 (.16)	.073
21 The ability of e-mail applications to automatically respond to all incoming messages with a return message specified by the recipient (e.g., "I am out of town this week.") is: (currently available; likely to be available soon; not technologically feasible	1.64 (.32)	-.96 (.21)	.13 (.09)	.308
Internet				
22 Which of the following is not a Web browser? (Mosaic; Yahoo; Internet Explorer; Netscape)	2.28 (.40)	.38 (.08)	.12 (.03)	.493
23 A file with a record of web site activity is called a(n): (hot file; click file; cookie file; active file)	.87 (.13)	-1.38 (.43)	.17 (.13)	.101
24 When you visit a website for the first time, which of the following things does the site know about you? (check all that apply) (E-mail address; The previously visited website; Social security number; The length of time you view each page on their website	2.45 (.64)	4.53 (.71)	.01 (.00)	.000
Information Search				
25 Which of the following online information sources is not free? (Lexis/Nexis; U.S. Census Bureau; Yahoo; U.S.A. Today)	1.29 (.90)	1.31 (1.17)	.52 (.15)	.034
26 Government regulation currently prohibits search engines from accepting payments in return for higher placement in the order of search results. (True; False)	1.24 (.69)	.83 (.49)	.31 (.11)	.097
27 Websites that assemble price comparison tables for products offered by multiple Internet stores are: (currently available on the Internet; will soon be available on the Internet; are not technologically feasible)	1.39 (.53)	-1.76 (.68)	.35 (.22)	.113

Table 2

Summary Statistics by Subscale for WVTM data

	Self-Report Mean	**Observed % Correct**	**Correlation**
E-mail	2.64	80.4	0.11
Internet	2.51	41.5	0.22
Information Search	2.47	70.5	0.10
Printing	2.39	66.0	0.15
Word Processing	2.01	72.9	0.32
Terminology	1.81	74.7	0.38
File Management	1.68	70.2	0.37
Spreadsheet	1.46	33.5	0.46
Data Base	1.35	55.0	0.17

Table 3

Antecedents of Computer Proficiency

Independent Variables	Tested Proficiency		Theta Hat		Self Rated Proficiency		Calibration=(Self-Tested)^2	
	Coefficient	t-statistic	Coefficient	t-statistic	Coefficient	t-statistic	Coefficient	t-statistic
Technology Familiarity	1.16	11.45	0.27	12.86	2.10	16.72	0.45	11.08
Years Online	0.77	10.59	0.16	10.97	0.72	7.98	0.17	5.67
Education	0.48	6.52	0.10	6.53	0.29	3.19	0.06	2.14
Technical	1.30	4.06	0.29	4.38	1.35	3.39	-0.01	-0.08
Age	-0.03	-3.93	-0.01	-4.51	-0.02	-2.02	0.00	-0.41
Hours Online/week	0.16	3.37	0.03	3.26	0.11	1.97	0.06	3.33
Student	1.22	3.08	0.29	3.58	1.53	3.14	0.37	2.36
Desire for New Electronic Gadgets	-0.30	-3.06	-0.06	-2.97	0.35	2.88	0.07	1.80
African American	-1.28	-2.13	-0.37	-2.96	0.79	1.06	0.34	1.40
Income	0.06	2.12	0.01	1.91	-0.01	-0.34	-0.02	-2.23
Caucasian	0.55	1.71	0.08	1.27	0.35	0.89	0.18	1.39
# Emails/week	0.08	1.62	0.02	1.67	0.25	4.05	0.03	1.42
Retired	-0.50	-1.35	-0.11	-1.42	-0.96	-2.08	-0.14	-0.95
Gender	0.05	0.26	-0.01	-0.25	0.21	0.87	0.09	1.20
Professional	0.04	0.13	0.04	0.76	-0.22	-0.63	-0.29	-2.51
Intercept	11.37	20.84	-1.10	-9.87	12.38	18.33	-2.41	-10.96
	R^2=0.369		R^2=0.403		R^2=0.378		R^2=0.199	

Table 4

Consequences of Computer Proficiency

	Independent Variables				
	Tested Proficiency		Self-Rated Proficiency		
Dependent Variables	**Coefficient**	**t-statistic**	**Coefficient**	**t-statistic**	**R^2**
Number of online purchases in the last 6 months	0.217	8.12	0.054	2.51	.076
Concern with the security of online transactions	-0.084	-6.30	0.008	<1	.031
Only buy products on the Internet from known and trusted companies	-0.028	-4.12	0.002	<1	.014
Concern about 3rd party monitoring or intercepting of online activity	-0.049	-3.72	-0.004	<-1	.013
Number of online purchases in the last 6 months as the result of clicking on a banner advertisement	-0.006	-1.17	0.008	2.04	.003