Cyberspace Representations and Choice

A. Norman, A. Agarwal, Y. Beyene, A. Dalal, M. DeGroot, C. Govil, J. Hart, N. Hite, K. Hood, M. Jindal, G. El Kadiri, S. Keyburn, M. Kulkarni, M. Rajan, J. Sanghai, P. Shah, V. Shah, Y. Sivakumaran, E. Sublett, A. Surratt, J. Sussman, M. Tang, C. Tillmanns, and F. Zahradnic

October 13, 2003

Abstract

Consumer search in cyberspace is more efficient than searching in physical space. Searching requires much less time and resources when going from site to site than from store to store. Also, cyberspace sites provide tools for consumers, such as tables listing attributes for side-by-side comparison and search software not found in physical space. But, in cyberspace, sensory perception is limited to sight and sound; for example, a consumer cannot try on a pair of pants to see if they fit. The extent to which consumers will switch their searches to cyberspace is directly related to the ability of entrepreneurs to develop alternative cyberspace representations of objects that compensate for the lack of sensory perception. We first consider digital cameras and the low resolution of computer screens. Next we develop an alternative procedure for evaluating pens. Then we consider finding pants that fit in cyberspace. We predict that, over time, consumer search will increasingly switch to cyberspace.

Send correspondence to:

Alfred Lorn Norman
Department of Economics
The University of Texas at Austin
Austin, TX USA 78712
norman@eco.utexas.edu

1 Introduction

This paper is one of a series of papers developing a model of a procedural consumer. The previous papers focused on ordering, Norman et al (2003), psychological decision rules, Norman et at (forthcoming), and budgeting, Norman et al (2001). These papers are also available in PDF format at http://www.eco.utexas.edu/Homepages/Faculty/Norman/research.htm. The current paper takes up the shift of consumer markets from physical- to cyberspace.

In evaluating an alternative for possible purchase, consumers forecast its performance in its intended use. As technological change accelerates, however, consumers' experience with previous products becomes less relevant to such predictions. Consequently, consumers demand data to augment their experience and allow successful forecasts. As presented in Section 2, the information value of such data is inversely related to its processing cost and positively related to its reliability and its discriminative capacity. Consumers' forecasts are based on a representation of the alternative, as, even in stores, modern packaging limits consumers' ability to physically interact with alternatives being considered. Is it more efficient for consumers to process these representations in physical- or cyberspace?

In cyberspace, consumers can employ tools to reduce the processing cost of analyzing data, but they lack sensory perceptions–such as touch–frequently used to evaluate products. For cyberspace to become a more efficient environment than physical space for evaluating alternatives, new representations must be developed to compensate for the inability to represent several types of sensory data. In Section 3, we consider the representation of digital cameras, pens, and women's pants in cyberspace. We consider whether subjects will rank the pictures taken by digital cameras in the same relative order, whether displayed on a computer screen or printed on photo paper by quality printers. In order to evaluate pens in cyberspace, consumers can be supplied pictures of the ink track to compensate for their inability to write with the pen to determine the smoothness of writing.

Women in physical space try on several pairs of pants to find a pair that fits; in cyberspace, women can learn to make good decisions using more detailed measurements, but perhaps a better long term solution is to provide women with a laser generated profile of their bodies. We also consider other statistics that could be employed in cyberspace to represent alternatives.

In Section 4, we discuss how cyberspace sites can provide consumers with useful tools to evaluate alternatives. Software running on the site server can use the polynomial processing ability of the computer to enhance the consumer's ability to analyze alternatives. The simplest case is the ability to create attribute tables on demand. Articles in trade magazines can also provide tables listing attributes side-by-side, but in cyberspace these tables can be constructed on the fly using only the consumer's desired alternatives. Another development in its infancy is the creation of artificial intelligence search procedures to find the desired alternative. We also consider future tool improvements.

Over time, consumer searches will increasingly take place in cyberspace, causing a significant reorganization of urban economies.

2 Cyberspace Representations

In order to select their preferred item, consumers need to forecast the future performance of the alternatives being considered. For inexpensive, frequently purchased goods, a consumer can learn to forecast performance through trial-and-error; one interpretation of the traditional utility model of the consumer is based on an adult who has learned her preferences through experience so there is no question of how a consumer forecasts performance. This method of gaining experience is increasingly infeasible, a consequence of an accelerating rate of technological change. Experience with previous products is no longer as accurate a guide to forecasting the performance of those current alternatives in the marketplace that are purchased infrequently or are undergoing a rapid rate of technological change. For most products, the ability to perform a test drive is limited. In such cases, consumers seek substitutes for personal experience to improve their ability to forecast future performance.

Consumers generally seek the experience of others. A consumer interested in buying a rocket to place a personal satellite into orbit could employ simulation programs that would accurately evaluate the performance of alternative rockets, but the vast majority of consumer products are not covered by models that will accurately simulate the performance of an alternative. In such situations, the best model available is others' experiences. One such category is current users, including people known to the consumer and anonymous user opinions collected by an Internet site; a second is the experience of experts who evaluated the alternatives through various tests.

A consumer must forecast the future performance of an alternative through its representation. This representation includes: data that clearly identifies the alternative; the sensory data obtained by physical interaction with the object; the opinions of current users–whether known or unknown; and the reports of experts based on tests of the alternatives. It is obvious that the consumer is evaluating a representation when she views alternatives through the Internet, but this also holds true for objects being evaluated in physical space. Physical interaction is frequently limited by modern packaging and the ability to test a product over its range of intended uses is impossible in most cases.

We start by considering the limits to the amount of data that a consumer could obtain to evaluate alternatives in the purchase of a vehicle, such as a car, an SUV, or a pickup. For this exercise, we shall consider all possible data obtained from all sources. The consumer could test drive the vehicle, study the vehicle brochure to examine all options, and look at the specification sheet detailing the values of the attributes of a vehicle, including the number of cylinders, displacement, and options present on the vehicle. The consumer might then examine the fuel economy and crash tests performed by various institutions, and finish her review by reading reviews performed by professional evaluators, such as Consumer Reports. This data is but a miniscule fraction of what could be provided on a vehicle. The potential buyer could be supplied with the detailed specifications of every component on the car, all the production details, every research and development report, and the details of all testing programs. The consumer could even be supplied pictures of the crystal structure of the metals in the vehicle. To the limit of current measurements, even the position of molecules on the surface of an object could be given. As Avogadro's number has magnitude of 10^{23}, a huge data file describing at least 10^{17} molecules on the surface could be obtained. Indeed, the limit of data that could possibly be supplied to a consumer is determined by the Heisenberg uncertainty principle. This exercise is intended to convince the reader that more data is not necessarily better than less; such a massive data file would be prohibitively expensive to obtain and would overwhelm the processing capabilities of the consumer. Indeed, it might actually lead to a poorer decision because the typical consumer could not interpret most of the data.

Thus, an important limitation to the consumer's demand for data is the processing cost. Processing data requires resources even in the absence of acquisition fees, but obtaining the specifications of alternatives can require examining several brochures or visiting several Internet sites. While a consumer aided by a computer can inexpensively process a great deal more data than a consumer unaided by a computer, the amount is still miniscule in comparison to the potential limit. In addition, the consumer faced with new technology or previously unencountered choices might have to acquire knowledge in order to interpret the relevant data. Consequently, the cost of data to the consumer should include acquisition fees, the cost of obtaining the knowledge to interpret the data, and the resources to acquire and process the data.

Another important aspect of the value of consumer data is its capacity to help the consumer discriminate between the alternatives' probable future performance. Akerloff (1970) was the first to raise this issue in his discussion of the used car market, where buyers lacked a means of judging the reliability of a used car. Spence (1973) expanded the discussion by pointing out the incentives of the market participant to provide signals that allowed the other party to make a judgement of future performance. In the case of used cars, auto dealers have developed a program of certification of the cars, where vehicles pass inspections and are deemed worthy of warrantees. This signal and others like it enable the consumer to discriminate among used cars. Signals provided by manufacturers include a specification list of product attributes and their values, such as the CPU and processor speed of a computer.

Third parties go further in providing consumers with data to discriminate among future performance of alternatives. Current users can communicate their experiences with a product, both good and bad; experts provide test-based reports evaluating the product and give advice as to what features are desirable for what purpose. These third party reports help consumers discriminate between products on the basis of the consumer's intended use of the product.

Since a great deal of data is available to consumers, the final factor considered in the information value of the data is the reliability of its source. Second parties have incentives to present the positive aspects of their products, while most third party information suppliers have a conflict of interest in that they are financed by advertising rather than by the prospective buyer. In summary, we define the information value of data as positively related to its ability to discriminate among alternatives and its reliability and inversely related to its processing cost.

Now let us use examples to illustrate this relationship. Our concept of information value measures the resources required to discriminate among alternatives. Basically, the lower the cost to discriminate, the higher the information value; the less reputable a data source, the more resources a decision maker must use in obtaining other opinions to confirm or refect the less reliable data. It is important to note that data sets consisting of as little as a single value of an attribute can have information value. Consider the value of the cycles/second of the CPU of a microprocessor in comparing personal computers: if all other factors of two personal computers are equal and the processor only differs by the cycles per second, this single value piece of data allows for discrimination between the two. But, the greater the variation in the architecture of the processor, the size of the cache, and the computer's RAM, the greater the bias in a single-valued discriminator. The consumer must consider multivalued data sets of the attributes that vary and the results of benchmark tests computing the same test problem. The consumer requires greater knowledge to interpret the data and must pay a higher processing cost to boot. Now consider a review of alternatives, say digital cameras, made by a reliable source: if the results are presented as a ranking, then this provides a single-valued data piece to discriminate. But even if the reviewer is reputable, the ranking may not correspond to that of the consumer if he has different criteria than the reviewer. Also, if the review influences the market by increasing demand for the top-rated camera, the consumer will pay a premium for the top-rated choice. We present these examples to illustrate the nature of the tradeoffs between resource cost to evaluate and the ability to discriminate.

The information value of a data set is a function of these three arguments, but in this paper we shall not try to explicitly construct such a function. We will rather consider the simpler problem of comparing two data sets A and B. If processing cost and reliability are equal for A and B, the set with the greater ability to discriminate among alternatives has the higher the information value; if reliability and ability to discriminate are equal for A and B, the set with the smaller processing cost has the higher information value; if processing cost and ability to discriminate are equal for A and B, the set whose supplier has the better reputation has the higher the information value.

Because consumers generally demand the same data, there are tremendous economies of scale in meeting their demand. In physical space, these economies of scale are realized in how-to books–such as "For Dummies" books–and in trade magazines–such as MacWorld–that review hardware and software. Currently, there are a growing number of sites in cyberspace specializing in providing consumers with data in order to evaluate alternatives. Internet sites currently cannot provide much of the data that consumers obtain from direct sensory interaction with the alternatives; consumers cannot try on a pair of pants over the Internet. The extent that consumer markets will shift to cyberspace is directly related to the ability of entrepreneurs to create new representations of objects that substitute for direct sensory interaction.

3 Compensating for Direct Sensory Perception

We consider three examples of the limits of sensory perception in cyberspace: digital cameras, pens, and women's pants. There are many Internet galleries of digital camera pictures that enable the consumer to compare the pictures of alternative cameras, but a computer screen is a low resolution device when compared to an inkjet or a laser printer; the issue to resolve is whether a consumer will make the same decision if she only views pictures on a computer screen instead of both viewing them on a computer screen and printing them with a quality printer. The second issue to consider is how to compensate consumers for their inability to perceive the smoothness of ink flow when evaluating pens on the Internet. The third issue is how the Internet can be used to find pants that fit well.

3.1 Digital Cameras

A very important aspect of judging alternative digital cameras is the differences in the quality of the cameras' pictures. We performed an experiment to investigate how much a subject's judgements of picture quality depends on whether the picture is viewed on a computer screen or photo paper. If the medium makes no difference, a consumer can make his decision entirely within cyberspace.

Rather than focusing on the pictures' subjects, we chose to focus on the underlying aspects of good pictures: color accuracy and saturation, white balance, and amount of detail. We obtained permission from Dave Etchell to use his copyrighted photo gallery posted at Imaging-Resource.com for the experiment; his website reviews cameras in part by taking pictures under uniform conditions.

This does not, however, address a major problem: there are variations in the color accuracy and saturation of every computer screen and printer. Even changing the brand of photo paper used in an inkjet printer can result in a perceptible shift in the picture's colors. Interested parties can purchase software to calibrate the appearance of colors for more than $100; alternatively, they could purchase a particular combination of computer screen and printer to resolve this difficulty. As few–if any–consumers would make such a purchase to evaluate digital cameras, we did not take this step.

Three pictures from each camera were selected for study:

1. An image of the GretagMacBeth ColorChecker, cropped from the "Davebox" picture. This picture provides a comparison for color accuracy and saturation. The actual ColorChecker is an 8.5" x 11" sheet with 24 equal-sized squares, each colored as written below:

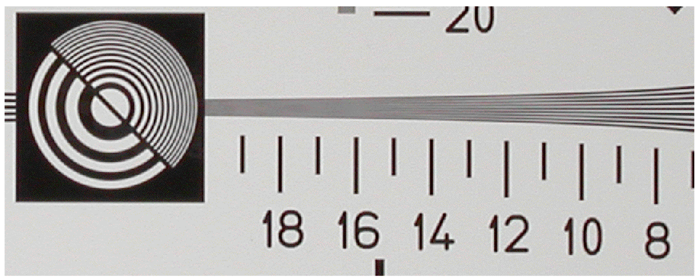
Table 1: GretagMacBeth ColorChecker [Text description rather than actual colors]

dark skin	light skin	blue sky	foliage	blue flower	bluish green
orange	purplish blue	moderate red	purple	yellow green	orange yellow
blue	green	red	yellow	magenta	cyan
white	neutral 8	neutral 6.5	neutral 5	neutral 3.5	black

2. An indoor portrait in incandescent lighting with no flash. A black and white version of this picture is below:

The woman is holding a bouquet of many-colored flowers in front of a white wall. The lighting in this picture is harsh, and many digital cameras take the picture with a yellow cast: this can make the background appear yellow, or even brown.

3. An image cropped from the resolution test picture. An example is shown below:



The colors in this picture do not matter.

3.1.1 Experimental Design

Four 4-megapixel cameras were selected for study: the Kodak DX4900, the Konica Digital Revio KD-400Z, the Kyocera FineCam S4, and the Minolta F100. These cameras represent a wide range of performance in the 4-megapixel category.

Subjects were paid a flat fee of $7 for about 10 minutes of rating three sets of pictures without a list of the camera names. The subjects first rated the pictures as displayed on a Sony CPD-G400 19" monitor; they then rated the three sets as printed on photo paper. The four ColorChecker and four white balance pictures were printed by Precision Photo on a Fuji Laser Frontier 390 printer. The four detail pictures were printed on a HP PSC 2175 printer. For each of the six sets of pictures, the subject saw all four of the pictures for the respective set simultaneously. For the ColorChecker and the white balance pictures, the corresponding Canon G2 picture was placed in the center as a reference. The pictures were sized so that they would appear on the screen simultaneously, and printed in a size so that the ColorChecker and detail pictures would all fit on a 11"x14" sheet of paper and the white balance pictures would fit on a 22"x14" sheet. The location of the four pictures on the screen was different from their location on paper.

Before rating the ColorChecker, white balance, or detail pictures, the subject was given a short training program in order to teach him what to look for in the respective pictures. For the ColorChecker and white balance pictures, the subject was asked to consider the Canon G2 picture as the standard and rate the other pictures as deviations thereof, with the picture with the greatest deviation given the lowest number > 0 and the picture with the smallest deviation the highest number < 9. The subjects were asked to consider all four pictures simultaneously so that their ratings would be consistent. For the detail pictures, the subjects were asked to rate the picture a 10 if they could see a separation between the lines at point 14, and less than that if the separation became blurred to the right of 14. They were asked to rate the pictures from 1-10 based on how far to the left they could see separation among the lines.

The experimental setup can be viewed at

http://www.eco.utexas.edu/Homepages/Faculty/Norman/CameraExpF03/cnn.htm.

3.1.2 Results

Below are three tables of mean ratings for the ColorCheker, white balance, and detail pictures of the cameras:

Table 2: ColorChecker: Color Accuracy and Saturation Ratings

Camera	No. Subjects	Screen Mean	Screen Std Error	Paper Mean	Paper Std Error
Minolta F100	49	6.63	0.15	7.24	0.14
Kodak DX4900	49	6.57	0.16	6.65	0.20
Konica Kd-400Z	49	6.43	0.18	7.27	0.15
Kyocera S4	49	2.78	0.16	3.00	0.17

Notice that the performance of the Kodak DX4900 in taking the ColorChecker picture is worse on paper than on the screen relative to the Minolta and Konica, and that the relative performance of the Kyocera S4 in taking the white balance picture is better on the screen than on paper. The relative positions of the cameras concerning the detail picture are the same on screen and paper.

Table 3: White Balance

Camera	No. Subjects	Screen Mean	Screen Std Error	Paper Mean	Paper Std Error
Kyocera S4	49	6.61	0.22	6.00	0.25
Kodak DX4900	49	6.14	0.14	6.73	0.17
Minolta F100	49	5.43	0.17	6.16	0.17
Konica Kd-400Z	49	3.51	0.14	4.14	0.17

Table 4: Detail

Camera	No. Subjects	Screen Mean	Screen Std Error	Paper Mean	Paper Std Error
Kodak DX4900	49	6.71	0.28	6.86	0.27
Minolta F100	49	6.63	0.26	6.76	0.26
Konica Kd-400Z	49	6.61	0.22	6.73	0.23
Kyocera S4	49	4.63	0.29	5.10	0.27

Are these declines statistically significant? We performed three 1x4 ANOVAs on the screen data and three 1x4 ANOVAs on the paper data for the three types of pictures, and performed a Duncan test to see how the cameras were grouped in each case. As one might expect, the hypothesis that the means were equal was rejected in all cases with α error of less than 0.0001 in each case.

Consider first the ColorChecker pictures. For the screen data, the Duncan test grouped the Minolta, Kodak, and Konica in the top group, and the Kyocera in the bottom group. For the paper data, the Duncan test grouped the Minolta and Konica in the top group, the Kodak in the second group, and the Kyocera in the bottom group. Thus, the experiment indicates the decline in Kodak DX4900 performance from screen to paper is statistically significant. The Kodak ColorChecker picture on the screen looks well saturated, but on paper the Kodak ColorChecker picture looks muddy.

Now consider the white balance pictures. For the screen data, the Duncan test grouped each camera into its own group, with the Kyocera first, the Kodak second, the Minolta third, and the Konica fourth. For the paper data, the groups changed so that the Kodak was first, the Minolta and

Kyocera second, and the Konica third. On the screen, the background wall in the Kyocera picture looked near white, but on paper the background wall looked a dirty grey; conversely, the Kodak picture looked much more pleasing on paper than it did on screen. We also compared the white balance pictures for the Kyocera, Kodak, and Canon from 4 printers: a Fuji Laser Frontier 390, a Sony Spectra Pix Digital Color Printer UP-D70A, an HP PSC 2175, and a Canon i950. Because the Kodak and Kyocera pictures looked best on different printers; it is possible that camera ratings would shift according to which printer was used.

Finally, consider the detail picture. For both the screen and paper data, the Duncan test gave the same result: Kodak, Minolta, and Konica were in the first group and the Kyocera was in the second group.

3.1.3 Interpretation

In order to place the 6 ColorChecker pictures on the computer screen at the same time, these pictures had to be reduced in size in Photoshop. In terms of accuracy and saturation, the Kodak Colorchecker picture does not appear to differ when reduced in size, implying that the effect of a shift from screen to printer varies among ColorChecker pictures. Examining the printed Kyocera white balance picture on screen in actual-pixels mode in Photoshop showed that the background looked much closer to white onscreen than it did when printed. Again, this indicates that the effect of a shift from screen to printer on white balance pictures varies among cameras, while in the case of detail the color shifts were not a factor in the subjects' judgements and the ratings were invariant between screen and paper.

Consumers can evaluate cameras using Internet picture galleries, but they should weigh their camera ratings by viewing the pictures on the screen and by examining printed pictures. Their weighing should reflect how much the consumer plans to view the camera pictures on screen versus on paper.

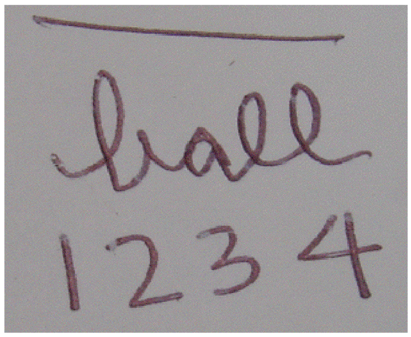
3.2 Pen Experiment

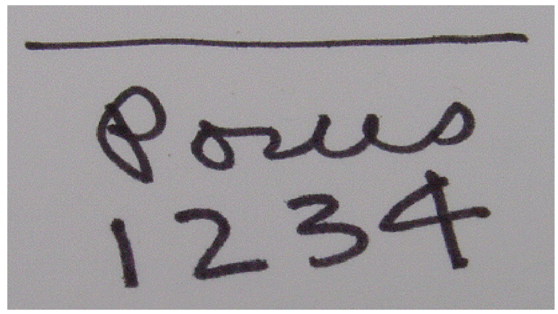
Most students consider how smoothly a pen writes to be an important factor in comparing two pens. The ability for subjects to successfully compare two pens over the Internet depends on creating new information that can substitute for the experience of writing with each pen.

3.2.1 Training Program

We created a training program based on pictures of writing samples of various pens. When first attempting to create this sort of substitute data, we focused on an opinion poll that would provide subjects with ratings on how smoothly each pen in the experiment writes. We found that subjects' opinions on smoothness were highly subjective; there was little agreement among polled students. The second attempt focused on training subjects to recognize the four types of pens–ballpoint, rollerball, gel tip, and porous tip–used in the experiment by writing with sample pens. Simply training the subjects to recognize the four types of pens did not appear to make a difference, so we reoriented the training program towards teaching subjects to carefully examine writing samples of the pens after learning to recognize these differences using the above image.

The type of ink constitutes the major difference among ballpoints, rollerballs, and gel pens. In a ballpoint pen, the ink is viscous and the pen tends to skip; in a rollerball pen, the ink is very fluid and the ink tends to bleed; in gel ink pens, the ink is less viscous than a ballpoint but less fluid than a rollerball, and the pen neither skips nor bleeds; ink from a porous pen, a pen with a felt or plastic tip, tends to bleed and also has a greater tendency towards a thick line than do other pens. Subjects in the experiment were trained to identify the four types of pens by observing skipping, bleeding, and the thickness of line. After examining the picture above, subjects were then shown a picture of four samples of writing and asked to identify the type of pen associated with each sample. Ballpoint and porous pens are easy to distinguish, whereas rollerball and gel pens are less distinctive, as can been seen by the two writing samples below:





Subjects in the training program were then given one picture of each type of pen, and told to rank them in order of preference. There were three pictures for each pen: a picture of the whole pen, a closeup of the pen point or tip, and a picture of a writing sample (such as shown above). All the data for the four pens were shown simultaneously in a frame, so that the subjects could rank them more effectively. After ranking the pens based on the pictures, the subjects were given each pen and told to rank them again after writing with each of them. This round was then repeated with a new set of four pens, again one of each type.

3.2.2 Experimental Design

The effectiveness of the training program was tested in a 1x4 ANOVA design.

Subjects were given pairs of pens and first asked how much prior knowledge they had of each pen: "None," "Some," or "A Lot." The subjects were then asked to compare the two pens based on the following data sets:

1. OfficeMax: The information provided on the OfficeMax.com site. The picture of the pen is too small and its description too vague to have much information value. For example, the OfficeMax description of the Sanford Uni-Gel Grip is "The Uni-Gel Grip pen combines gel ink technology with extreme comfort. The rubber grip provides writing comfort and control. Gel ink provides smooth writing and rich vibrant colors."

2. Better Pictures: Images were taken with a Sony CyberShot 2-Megapixel digital camera to obtain an image of the pen much larger than that on the OfficeMax website; a closeup picture of the pen point was also included with the assumption that subjects would associate the point with smoothness from experience.

3. Writing Sample: The same as “Better Pictures,” but with the addition of a writing sample, such as the two samples shown previously, to be interpreted with the training program described above.

4. Actual Writing: For each comparison, subjects were given the two pens and asked to write with them before making their choice.

In their comparison of pen A and pen B, the subjects were asked whether they strongly preferred pen A, SP(A); preferred pen A, P(A); were indifferent between the two, I; preferred pen B, P(B); or strongly preferred pen B, SP(B). Then the subjects were asked to again compare the two pens based on actual writing. They sequentially processed 7 pairs of pens in this fashion.

The four treatments were:

- OfficeMax followed by Actual Writing
- Better Pictures followed by Actual Writing
- Writing Sample followed by Actual Writing
- Actual Writing followed by Actual Writing. All of the pairs of pens were processed once and were then reprocessed, using the same brand but in a different order.

To obtain a statistic measuring the regret of the choice based on the first information structure, we defined the reversal rate between the first and the second information structure in the following table:

Table 5: Definition of Reversal Rate between Pen A and Pen B

First Rank	Second Rank	Reversal Rate
SP(A) or P(A)	SP(A) or P(A)	0
SP(B) or P(B)	SP(B) or P(B)	0
SP(A) or P(A) or SP(B) or P(B)	Indifferent	0.5
SP(A) or P(A)	SP(B) or P(B)	1
SP(B) or P(B)	SP(A) or P(A)	1

There were 40 subjects drawn from the undergraduate population of the University of Texas at Austin, and 10 subjects were assigned to each treatment. The ANOVA was performed on the mean reversal rate for each treatment.

3.2.3 Incentives

Each subject was paid a flat $5 and was told he would receive two pens that he indicated that he preferred or strongly preferred. Each subject was told these two would come from one pair based on the first information structure and one pair based on the second information structure, with the pairs being chosen at random.

The experiment can be viewed at http://www.eco.utexas.edu/Homepages/Faculty/Norman/penexp.html.

3.2.4 Results

Table 6: Analysis of Variance Procedure

Source	DF	Anova SS	Mean Square	F Value	Pr > F
Treatment	3	20.31875000	6.77291667	8.08	0.0003

The ANOVA rejects the hypothesis that the mean reversal rate for the four treatments is the same.

We also considered the Duncan test to see how the four treatments were grouped. In the table below, the means with the same letter are not significantly different.

Our experiment demonstrates that the information supplied by OfficeMax leads to decisions that

Table 7: Duncan Test

Duncan Grouping	Mean	N	GR
A	2.9500	10	1
A			
A	2.9500	10	2
B	1.5500	10	3
B			
B	1.5000	10	4

are only slightly better than choices made at random (mean = 3.5); with our training program, subjects' performance was on a par with decisions made after actually writing with the pens. Although the writing sample provides data with high information value so that subjects can discriminate, with a much larger sample it may well be that the writing sample is not quite as good as actual writing.

3.3 Woman's pants

When buying a pair of pants in a physical store, a woman generally must try on several pairs before finding one that fits. In a survey of 62 UT college women, we found that 61% tried on 3-5 pairs before making their purchase and 21% tried on 6-10 pairs; in the same survey 95% said that they would buy pants online if they could be assured of a perfect fit (guaranteed satisfaction). Difficulties in finding a good fit arise from the differing size systems used by manufacturers. For example, one of authors found that a size 4 from the Gap is most similar to a size 6 pair from Calvin Klein and a size 2 from Abercrombie. Because manufacturers' sizes are not standardized, women must rely on their sense of touch to find pants that fit. To increase the percentage of pants ordered online, vendors in cyberspace must devise a new system to substitute for touch.

Archetype Solutions, archetypesolutions.com, came up with a system of measurements that was used successfully by Land's End, landsend.com, to substitute for touch in creating made-to-order jeans and chinos. They expected 10% of their sales to be made-to-order, but the actual percentage of sales was much higher while their return rate was lower than the industry average of 6%. Land's

End's made-to-order operation was profitable because of supply chain management; the greater the percentage of made-to-order sales, the smaller the inventory Land's End must carry.

The women who buy clothes from Land's End are generally older than college-aged women. To see if the Archetype Solutions procedure would work with college-aged women, two of the women in our research group ordered a pair of chinos online through Bobs Stores online, bobstores.com, that recently adopted the Archetype Solutions procedure. This website offers the customer a choice of the style of chinos they want (regular or relaxed, cuffed, boot cut or straight, one or two pockets, etc.) before asking the customer to input a series of 10 body measurements and answer 6 questions on body proportions. After submitting these measurements, both young women received poor fits; both pants were too loose in the hip and thigh areas. We felt that this fit might be more popular among the population of older women, but for the junior/misses/petite customers, this more relaxed fit was not preferred. The website did not ask the age of the customer, which might be an area of interest in assessing the type of fit a customer would like. An additional problem was that, while one girl received her pants within a span of 3 weeks, the other received hers in 8 weeks. The reason for the delay was that a pattern was not found for the second womans measurements. Given a 6% reject rate, the probability of getting two pants too loose by chance is 0.4%. Bob's shuttered its made to order operation on 29 Aug 03.

How can the system of measurements be improved? Archetype Solutions might improve the fit generated by their case-based reasoning software by taking more measurements; however, in talking with a representative of the firm we learned that measurements supplied by customers are frequently inaccurate. Another approach, taken by AmericanFit, americanfitclothing.com, is to provide the consumer with detailed instructions on how to measure a pair of pants that fit in 10 defined locations so that americanfitclothing.com can replicate the dimensions. This solves the problem if the woman has a pair of pants that she considers is a good fit.

An alternative approach is to obtain a complete body image using a laser scanner. Levi Strauss has laser scanners in several stores that create a complete body image of the customer, but this approach requires the customer to strip down to their underclothes. An alternative laser scanner, developed by Intellifit, made4me.com, scans the customer's body in 10 seconds and allows the customer to remain fully clothed. Such a device in a retail store could quickly direct customers to pants that fit the first time. If the customer could further keep the measurements on electronic media, the customer could use them to buy made-to-order pants online. This approach appears to have the most long term potential.

3.4 Other Examples

In this section, we discuss some examples of products for which alternative cyberspace representations could improve the efficiency of consumers.

In buying thin-skinned fruit in a physical grocery store, a consumer uses touch, smell, and sight to judge the ripeness of the alternatives. In buying fruit online, a consumer can neither touch nor smell a piece of fruit to determine its ripeness. A new approach to judge the ripeness of fruit online has been made possible by Cemagref's (2000) development of techniques to measure its sugar content. While Cemagref is focusing on the market for sorting freshly picked fruit, their technology could someday be employed by online grocery sites.

One important factor in buying clothing is the feel of the material. To develop an alternate, objective measurement to judge the feel of the material, online clothing stores could employ the measurements developed by the fabric and clothing industry that substitute objective statistics for panels of professional testers. Product testers in the fabric industry routinely use tests such as the ring test and the slot test to obtain statistical measures relevant to texture, weight, feel, and other characteristics, Dent (2000). If consumers could be trained to interpret these data they would be able to judge the feel of clothing they purchased online, because such tests in the fabric and clothing industry correlate well with the judgement of professional testers.

The present inability of cyberspace to communicate smells to consumers affects the sale of a variety of online goods such as perfume. However, this barrier is slowly being overcome as well. DigiScent, digiscents.com, has been developing hardware and software that can be used to communicate smell online. Their technology uses replaceable scent cartridges to create different smells from a scent palette depending on the instructions received from the software. Products like this improve the efficacy of the online sale of goods for which smell is important.

As Internet consumer markets develop, entrepreneurs will have incentives to develop new approaches to represent alternatives so that consumers can forecast future performance.

4 Cyberspace Advantage

One area where cyberspace dominates physical space is in the provision of tools to consumers. Let us start with a computer search through cyberspace. Suppose a consumer knows what she wants to buy: she can start with a search engine like google.com and input the product name and the word 'price.' Google will return several sites, such as pricegrabber.com, that sell the product through linked merchants. Going to any of these sites, the consumer can immediately obtain a price distribution of the product and an estimate–frequently given as a number of stars–of the reliability of the seller. Obtaining such information in physical space from a large number of sellers would be prohibitively expensive because of travel costs or time delays in reaching a salesperson by phone. For new technology, consumers frequently want expert reviews. While such reviews are published in magazines, such as MacWorld, trying to remember the number of the appropriate issue and physically searching through indices can be a frustrating experience in physical space. Again, if a consumer goes to google.com and enters the product name and the word 'review,' she obtains a list of sites that have reviewed the product. At such a site, a consumer can click through a hierarchal structure to read the reviews of interest. The polynomial processing capabilities of computers connected in cyberspace make consumers much more efficient in obtaining price and product information than they are in physical space.

In addition, consumers visiting sites on the Internet can create tables of the attribute values of products of interest. In review articles in media magazines, the authors frequently include tables that show the discussed products' attributes side-by-side, but these tables are immobile. At third party cyberspace sites focusing on the same products, consumers can use software to create desired tables of the alternatives of interest. These tables can include the products' attributes, links to reviews, and current prices. In physical space it would be too expensive for consumers to create such tables by lining up the manufacturers' brochures that the consumer would have to collect from widely dispersed, alternative sources, such as dealers' showrooms. The creation of such tables allows the viewer to more easily ascertain just what attributes each alternative has, and such tables are much more accurate than remembering the information seen in alternative brochures. A consumer can create tables of attribute values for a wide range of products–including automobiles, computers and consumer electronics, household appliances, and software–at general sites, such as pricegrabber.com, dealtime.com, epinions.com and specialized sites such as autotrader.com and dpreview.com

Consumers at third party data sites can also employ search algorithms to reduce a large number of alternatives to a small number for more detailed study. Most of these algorithms use noncompensatory rules in that the consumer specifies a range of attribute values and the algorithm finds those alternatives that satisfy the criteria. In most cases, the consumer can then create a table describing the subset of selected alternatives. A smaller number of algorithms, such as those at ActiveBuyersGuide.com, rely almost exclusively on compensatory rules that weigh tradeoffs between attribute values. The weights used by the algorithm are set by asking the consumer a series of questions which enable the consumer to consider the consequences of a variety of criteria. Such exercises help the consumer decide what criteria are really important to him.

The construction of tables and search algorithms is in its infancy. The issue with tables is how much data to display. Too much data can overwhelm the consumer's processing capability. Tables for consumers should be organized by the current trend in organizing file structures by operating systems such as Windows or MacOS. The user should be able to expand or collapse the file structure as needed. Because consumers have varying levels of interest and prior knowledge, collapsible and expandable tables would be of greater use than the current one-dimensional tables. A great deal of research is needed to make search algorithms at third party sites more useful. At present, economists know almost nothing of the search procedures actually used by consumers.

5 Conclusion

The shift of consumer markets to cyberspace depends on the ability of entrepreneurs to develop new representations of products to compensate for the lack of sensory perception normally obtainable through direct physical interaction. We have provided examples to demonstrate that such substitutions are possible. Currently, data tables on the Internet are more useful than data tables in magazines and brochures. As research progresses, the quality of Internet search algorithms should also improve.

Another factor in the consumer's switch from physical markets to cybermarkets is the value of time to household members. The premium on time has increased as in increasing number of households all adults work. The fact that shopping on the Internet is much more time efficient than shopping in physical space will tend to accelerate the shift in markets.

Finally, small parcels shipped from warehouses can be aggregated through the use of logistics software. At the local level, a delivery truck delivering packages to households is much more efficient than a member of each household driving a large SUV back and forth to a mall to purchase a small number of items.

References

G. Akerloff, The Market for “LEMONS”: Quality Uncertainty and the Market Mechanism, *Quarterly Journal of Economics*, Vol. 89, (1970) 488-500

R. Dent, , An Analysis of Fabric ‘Hand’ and ‘Feel’, *International Nonwovens Journal*, <http://www.inda.org/subscrip/inj00_1/p11.html>(2000)

Cemagref, Automating the measurement of sugar content in fruit using a non-destructive method. <http://www.cemagref.fr/English/ex/agri-food/ex23.htm> (2002)

A. Norman, A. Ahmed, J. Chou, A. Dalal, K. Fortson, M. Jindal, C. Kurz, H. Lee, K. Payne, R. Rando, K. Sheppard, E. Sublett, and I. White, , On the Computational Complexity of Consumer Decision Rules, *Computational Economics* (forthcoming)

A. Norman, A. Ahmed, J. Chou, K. Fortson, C. Kurz, H. Lee, L. Linden, K. Meythaler, R. Rando, K. Sheppard, N. Tantzen, I. White and M. Ziegler, *An Ordering Experiment, Journal of Economic Behavior and Organization*, Vol 50 (2003), 249-262

A. Norman, J Chou, M Chowdhury, A Dalal, K Fortson, M. Jindal, K. Payne, and M Rajan,, Two-Stage Budgeting: A difficult problem, *Computational Economics*, 18 (2001) 259-271

M. Spense , Job Market Signaling, *Quarterly Journal of Economics* 87 (3) (1973) 355-374