Using Technology to Make New Assessment Instruments

G. Donald Allen, Ph. D.
Department of Mathematics
Texas A&M University
College Station, TX 77843
dallen@math.tamu.edu

Dianne S. Goldsby, Ph. D.
Department of Teaching, Learning, and Culture
Texas A&M University
College Station, TX 77843
dgoldsby@neo.tamu.edu

Abstract. The use of visual analog scales (VAS) in survey research has been rare, partly due to operational difficulties. Moreover, without mechanical devices, they have been difficult to implement. However Internet based surveys permit the use of continuous input devices such as slider bars, making VAS more feasible. The authors have devised a new slider tool and conducted an experiment to explore the utility of a VAS in a web survey, comparing it to radio button input. The emphasis here is on the technology of the instrument itself and some of the associated mathematics and results.

Surveys, Animation Technology, and the Problem

From first steps, we consider radio button surveys. On the Internet, these are analogs of survey instruments we have all used originally in paper form. The evaluation form for this conference is for the most part one of these survey types. Now we often see them posted to the Internet. Few of us have not been solicited to complete some such survey. The valuations of these radio buttons frequently range from "Strongly Disagree" through "Strongly Agree." The numbers of buttons are usually two, three, or five. Occasionally, there are other numbers of buttons. Such surveys were are just about the most sophisticated preference instruments possible with paper, and until recently on the internet. We are excluding here short answer or essay type questions, where the richness of the response sets often precludes simple analysis.

Recently, we have seen the advent of Internet based slider bars or visual analog scales as survey instruments. Such tools may help in the measurement of concepts using bipolar ratings, which can be hard to convey with numbered or labeled scales. For example, we are asking respondents to judge the extent to which certain behaviors or attributes are determined between two extremes such as environment vs. genetic or punishment vs. persuasion. These have the advantage of allowing the respondent to precisely place the slider position where they view it should be.

Technically, these devices work by JavaScript, Java, or Flash extracting the mouse position and rendering it as a numerical value. Only recently has research been conducted to compare slider bar information with radio button information [Arnau, et. al., 2001, Singer, et. al, 2004, and Couper, et. al, 2006, Torrance, et. al., 2001]. Results are not overwhelmingly positive, but the consensus is that there is a useful role for VASs in the process of measuring preferences.

The focus of this paper is on aspects of a two dimensional slider, i.e. a slider with two degrees of freedom of movement. In particular, we have developed a triangular slider, wherein respondents can compare continuously three competing attributes. For example, on cuisine the comparison may be between American, Italian, and Mexican. The term "triad" is not new in education and psychology literature. In most cases, a triad is a set of ranked responses between three items, determined by asking for a ranking between successive pairs of the items. An example may be the ranking preferences between the three colors, red, blue, and green. What can happen, and apparently does happen is that the ranking is intransitive. In the colors example this may be reflected by the subject preferring red to blue, blue to green, and then green to red. On these so-called stimulus circular triads there is a body of literature typified by Hendel (1977).

(Palmer and Pella 1975) performed a concordance analysis of preference among three items by asking the survey-taker to select the most preferred and the least preferred item of three. This automatically ranked the three items in preference. Their analysis proceeded to validate this triad type of comparison with the alternative of pair wise comparisons. Thus, the two rank orders were compared. (Paul, 2004) performed a triangular-type survey was done on paper with the triangle carved into subsets and values attributed to them. No algorithm was given for the subdivisions.

The slider, in the shape of a button, can be moved anywhere in the equilateral triangle. In Figure 1, we show a triangular slider with the button in two different position.

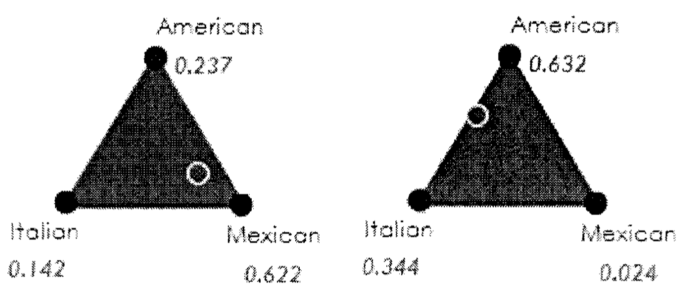

Figure 1

The values calculated are the coefficients of the button written as a convex combination of the three vertices. This effectively makes these values coordinate free. So, if v_1, v_2 and v_3 denote the vertices, and P is the position of the button, then

$$P = \lambda_1 v_1 + \lambda_2 v_2 + \lambda_3 v_3$$

A very important question is whether the selection of these coefficients in any way actually measures the true position of the respondent's preferences. What is important, therefore, is to design an experiment to make these conclusions. Another question is whether there are alternative measurements that may also accurately measure preferences.

The second question has an interesting resolution. An alternative geometry would be to measure the value of one vertex to be proportional to its distance to the opposite side. What this suggests is the way the respondent chooses between three options in a continuous way is by some mechanism of avoidance of one, or two of the options in favor of the other two or one option, respectively. Assuming the triangle has a height one, it is easy to see that all three distances are less than one. Indeed, for an equilateral triangle of height one, it is a well known theorem that the sum of the distances of a given point P to each of the sides sum to one. Now, the distance approach has a certain visual intuition, whereas the convex combination seems to merely be convenient. However, the following relates the two ideas. The prove is easy, and it possibly a new result.

Theorem 1. Given an equilateral triangle with height one, with vertices v_1, v_2 and v_3 and P a point in the triangle. Let P be written as a convex combination of the vertices, $P = \lambda_1 v_1 + \lambda_2 v_2 + \lambda_3 v_3$. Let d_1, d_2, and d_3 be respectively the distances from the vertices v_1, v_2 and v_3 to the opposite side. The $d_i = \lambda_i$, for $i = 1,2,3$.

Note this is trivially true in the analogous case of a slider bar with values calibrated between 0 and 1. There is an interesting generalization of this theorem to any triangle, where it is necessary to use scale factors given by the heights.

Technology of the Triad Survey Instrument

The triad instrument was programmed in Action Script, the programming language native to Flash. It is fairly simple code. There are two main portions of the code. The first is the code that activates the mouse movement when the button is pushed and then converts the mouse (i.e. button) to the appropriate values. The second part, though minor, plays the important role of not allowing the button to escape the triangle with for example a sudden mouse movement. The code for the movement portion is:

```
onClipEvent (mouseDown) {
    if (Math.abs(_root._xmouse-this._x) < 5 && Math.abs(_root._ymouse-this._y) < 5) {
startDrag(this, true, _root.leftt, _root.topp, _root.leftt+100, _root.topp+80);
        _root.go = 1;
    } }
onClipEvent (mouseUp) {
    stopDrag();
    _root.go = 0; }
onClipEvent (enterFrame) {
    if(_root.go == 1){
        _root.xx = _root.x-_root.leftt;
        _root.yy = _root.y-_root.topp;
      if (_root.yy > 80) {
        _root.yy = 80;       }
      if (_root.yy < -2) {
        _root.yy = -2;       }
a1 = (80-_root.yy)/76;
```

```
a2 = (_root.xx-100+a1*50)/(-100);
a3 = 1.0-a1-a2;
a1 = Math.max(a1, 0.0);
a2 = Math.max(a2, 0.0);
a3 = Math.max(a3, 0.0);
sum = a1+a2+a3;
_root.l1 = Math.round(1000*a1/sum)/1000;
_root.l2 = Math.round(1000*a2/sum)/1000;
_root.l3 = Math.round(1000*a3/sum)/1000;
}
else if (_root.prestart == 0) {
        _root.l1 = .333;
        _root.l2 = .333;
        _root.l3 = .333;
        _root.prestart = 1;
}}
```

So, it is seen there is a check to see if the mouse is in a close neighborhood of the button when it is pressed. If so, it is made "draggable." When the mouse button is released the action stops and the position is reported.

There is some difficulty in extracting the resultant triple of values from the Flash code to the main HTML page. This is done through the LiveConnect. This was originally Netscape Navigator code to allow plug-ins to be scripted. The class could be called from JavaScript and even from other Java applets running within the page with the browser marshalling the calls between the various contexts. It does work with IE from 4.0 and on, but it doesn't work with the Mac version of IE. In any event the results are communicated from the Flash executable embedded in the HTML page, and then posted through the usual forms mechanism. For example, we used <form name="survey" method="post" action="preferences_results.php" onSubmit="return validate()">. A PHP script was use to file the data. The sample that generated with results above is located at http://www.math.tamu.edu/dallen/etc/cuisine_sample.html. The source code is available at
http://www.math.tamu.edu/dallen/etc/cuisine_sample.fla

Comparative Results

The original use of this survey was to measure perceptions on teaching of pre-service teachers [Allen, et. al. 2001]. However, interesting as the results were, it is important to validate the instrument by comparing triadic and radio button results of two identical surveys given to the same sample of people. Thus we attempted to demonstrate that the simple triadic geometry is interpreted essentially in the same manner by almost every test-taker. Moreover, it is desired to show all or most participants use the triad in essentially the same way. For example, if there is a choice between three items where it is known a priori what the expected selection should be, does the participant move the dot into that vertex of the triangle? Similarly, if the choice is expected to be somewhere between two of the vertices, does the participant move the dot to the edge of the triangle containing those vertices? Our methods of validation follow:

A. Formulate a simple preference survey with highly predictable answers.
B. Execute this survey in two ways, one using the triadic instrument and the other using standard radio buttons.
C. Compare the results.

We offered eight questions of which just a sample is given below. We wanted to pick very simple alternatives where we had a strong suspicion of what the results should be. Here are just four of the questions.

2. My favorite type of cuisine is (choices: American, Italian, Mexican)
3. During the fall my favorite sport is (choices: Football, Basketball, Baseball)
4. My least favorite insect is (choices: Fire ant, Ladybug, Roach)
5. My favorite snack foods tastes (choices: Salty, Sweet, Bitter)

For each question of the radio button survey, the chosen answer was given a value of 1 with the other two given a value 0. In the language of the convex combinations and the triangle, this is the numerical equivalent of the user placing the button exactly at the vertex. For the triadic survey, the sum of the values was one. It is convenient to refer to the sum of these values as generalized frequencies. A standard chi-squared analysis was performed on the results of both surveys. Our results established there is no statistical difference between the two instruments, and this is interpreted to mean our interpretation of the geometry of distance to a side is commensurate with true preference values.

References

1. Allen, G. Donald, Goldsby, Dianne S., Kelly, Larry. Pre-Service Teacher Perceptions of Mathematics/Science Teacher Preparation, paper presented at AACTE conference, Austin, TX, October 2005.
2. Arnau, R. C., Thompson, R. L., Cook, C. . (2001) . Do Different Response Formats Change the Latent Structure of Responses? An Empirical Investigation Using Taxometric Analysis . Educational and Psychological Measurement, 61, 1, pp. 23-44
3. Couper, M. P., Tourangeau, R., Conrad, F. G., Singer, E. . (2006) . Evaluating the Effectiveness of Visual Analog Scales: A Web Experiment . Social Science Computer Review, 24, 2, pp. 227-245
4. Hendel, D.D. (1977). Intransitivity on paired-comparisons instruments. The relationship of the total circular triad score to stimulus circular triads. Applied Psychological Measurement, 1(3), 403-11. EJ186179
5. Palmer, G.A. & Pella, M.O. (1975). Determination of the relative consistency and concordance of student science interest responses by utilizing paired and triad stimulus comparison inventories. Science Education, 59(1), 19-26.
6. Paul, Jody. http://www.jodypaul.com/ASSESS/CBAA.pdf September 2004.
7. Singer, E., Tourangeau, R., Conrad, F. G., Couper, M. P. . (2004) . Evaluating the Effects of Visual Analog Scales: A Web Experiment . RC33 6th International Conference on Social Science Methodology: Recent Developments and Applications in Social Research Methodology, 2004.
8. Torrance, George, Feeny, David, and Furlong, William, Visual Analog Scales: Do They Have a Role in the Measurement of Preferences for Health States?, Medical Decision Making, Vol. 21, No. 4, 329-334 (2001)