Review of reports on consumer testing results for HUD's Good Faith Estimate (GFE) form

Report by
Janice (Ginny) Redish, Ph.D.
Redish & Associates, Inc.
6820 Winterberry Lane
Bethesda, MD 20817

301 229 3039
ginny@redish.net
www.redish.net

September 2007

Review of reports on consumer testing results for HUD's Good Faith Estimate (GFE) form

Introduction

The Department of Housing and Urban Development (HUD) through HR Communications, Ltd. asked Redish & Associates, Inc. to provide an expert review of four documents. The documents comprise the final reports of several rounds of iterative development and testing of new forms for people buying homes. The iterative development and testing was done by the Kleimann Communication Group (KCG) of Washington, DC., and, therefore, the documents are also known as the Kleimann Reports.

What four reports does this review cover?

The four reports covered by this review are:

Testing New HUD Forms with American Homebuyers	final report (of rounds 1 & 2) March 7, 2003
Testing New HUD Forms with American Homebuyers	third round testing report October 8, 2003
Consumer Testing Results for HUD's Good Faith Estimate (GFE) Form	rounds 4 & 5 report, March 2004
Volume II, Appendices – Consumer Testing Results for HUD's Good Faith Estimate (GFE) Form	rounds 4 & 5 testing materials report, March 2004

Who reviewed these four documents?

Dr. Janice (Ginny) Redish, President of Redish & Associates, Inc., Bethesda, Maryland, reviewed the four documents. For more on Dr. Redish's qualifications, see the biography at the end of this report.

What questions were the focus of the review?

HUD asks these four questions for a review:

- Does the paper involve significant errors of fact, methodology, or analysis?

- Does the paper have significant omissions or irrelevant materials that could be deleted?

- Can any stylistic improvements be suggested that would enhance the paper's readability?

- Does the reviewer have other comments on the paper?

These questions frame the following review, which covers all four reports as "the paper."

Does the paper involve significant errors of fact, methodology, or analysis?

No. The four reports are based on excellent methodologies carefully carried out and on appropriate analyses of the data collected in each round of testing.

Through several projects, KCG used the state-of-the-art, user-centered design process of iterative design – development – evaluation to move ever closer to forms that homebuyers can understand and use. They focused the effort on meeting HUD's specific goals, which are clearly stated in the reports.

The reports primarily detail the methods used to evaluate successive drafts of the forms.

The first three rounds of evaluation

The first three rounds of evaluation reflect best practices in at least these ways:

- KCG recruited for diversity among the participants.

- They evaluated the forms in several locations, achieving geographic diversity.

- They included 15 people from each city in each round, giving them 45 people in each of rounds 1 and 2 and 60 people in round 3.

- They carefully planned and carried out studies in which they dealt appropriately with the concerns and constraints of such studies. For example, when asking participants to use two different forms, they gave half of the participants one form first and the other half the other form first. This is necessary to control for "practice effects" – that people carry over what they learn from working with the first document to working with the second document. The accepted research method for handling practice effects is to alternate which document people work with first. KCG did that.

- They triangulated (used multiple methods) to understand how the forms were working for people and what specific aspects might need improvement. Specifically, they

  – allowed participants time to look over and comment on the forms (undirected observations with think aloud; to observe users' spontaneous behaviors and to record users' spontaneous comments related to the form)

  – gave participants tasks to do with the forms (to observe specific behaviors with the forms and to have data on accuracy of finding and understanding information and of using that information to make informed decisions)

  – focused users on specific parts of the forms to gather data on aspects of greatest concern for HUD's goals

- They analyzed the data appropriately and reported it in relation to each of HUD's goals.

- They used the results to improve the forms and tested them again. (That iteration is a key component of a successful forms design methodology.)

The fourth and fifth rounds of evaluation

The last two rounds of design – development – evaluation were large-scale quantitative tests.

Rounds 4 and 5 were done in part because another federal agency, the Federal Trade Commission (FTC), had conducted research that suggested the Yield Spread Premium might not be helping homebuyers. FTC conducted quantitative testing of *excerpts* from the GFE forms. From that testing, FTC concluded that including the Yield Spread Premium (YSP) confused people and that there might be a bias against brokers in including the YSP.

In rounds 4 and 5, KCG gave participants *entire* GFE forms. Thus, the KCG rounds 4 and 5 tests were more realistic than the study conducted by FTC. Homebuyers will receive entire forms, not excerpts from the forms.

In each of these rounds, KCG did large-scale testing of 600 people recruited in 5 cities. They divided the participants into groups so that they could have direct comparisons of forms with and without the YSP sections and with various versions of the YSP section.

The groups were large enough (150 people per group) to allow analysis with confidence in the answers to the specific questions of the studies.

The forms that KCG used in the rounds 4 and 5 tests were filled in with relevant data so that participants in the study were using the GFE form in simulated but realistic situations. The forms were carefully set up so that in one set a loan offer from a broker cost less and in another set the loan offers cost the same. Thus, the situations were set up so that KCG would see if participants chose loans that indicated a bias against brokers.

All of these points show the care with which KCG developed and used the materials and the methodology. Both were sound.

When they found in round 4 that there were areas of misunderstanding, KCG changed the language and layout and tested again (round 5). The improvements in scores in round 5 show the value of the iterative design – develop – evaluate approach to creating useful forms.

With the changes tested in round 5, KCG was able to present a GFE with YSP disclosure to HUD that achieved very high accurate performance from study participants with no bias against brokers. The accuracy rates of being able to identify the less expensive loan were exceptionally high: 92% (broker loan cheaper), 93% (lender loan cheaper), 92% (loan costs the same). In usability studies, when setting expectations for success, the highest measurable usability goal is typically that 90% of participants will be able to do [x]. (Only in mission critical situations, such as airline pilots flying planes correctly, are usability goals set at 100%; and such performance almost always requires extensive, detailed, and expensive training to achieve.)

Does the paper have significant omissions or irrelevant materials that could be deleted?

No. The reports are well written and presented.

The reports begin with useful context-setting information so that the reader is brought into the ongoing story and understands the motivations for and goals of each round. The

results are presented in logical order related to the goals and to the comparisons being looked at. The reports are thorough, with appendices that give all the necessary materials for someone to understand exactly how people were recruited, what they were asked / told to do, and the forms they worked with.

Can any stylistic improvements be suggested that would enhance the paper's readability?

For the most part, the presentation is very clear.

I caught a few glitches and have a few suggestions for improvements. I give those in the next section.

Does the reviewer have other comments on the paper?

I have three comments.

Important note: I stress that these comments are all minor. None affects the reliability of the methodology and analysis nor the validity of the findings in these reports.

Comment 1: The name "first-time homebuyers" is not accurate for the group recruited as "people actively looking"

KCG recruited from two groups of homebuyers:

- people who have bought or refinanced a home (house, townhouse, or condominium) within the last two years

- people who have not bought within the last two years but who have actively looked within the last six months (defined as going to an "open house" of a house for sale, contacting a real estate agent, or pre-qualifying for a mortgage loan).

Recruiting within these two groups assured that participants would include

- people with recent experience of actually purchasing or refinancing a home, and

- highly motivated people without recent experience.

Once recruited, participants became part of the general pool and no further analysis was done based on their status.

All that is fine. I have no problems with it.

My concern is *only* with the way the two groups of recruited participants are named in the report. Those who have bought or refinanced in the last two years are called "experienced homebuyers." Fine. No problem.

Those who have not bought in the last two years who are now looking are called "first-time homebuyers." That is a problem. The recruiters did not ask whether these people had ever owned a home. Many current lookers are indeed first-time homebuyers. But current lookers may also be people who now own a home that they bought more than

two years ago and have not recently refinanced. They could be on their second, third, … nth home.

I recommend using names for these groups that more accurately reflect them. Perhaps:

- homebuyers with recent experience of loan documents

- current lookers without recent experience of loan documents

Because the distinction is needed only to describe recruiting (not to explain the evaluation techniques or to present the analysis), the long names should be acceptable.

Comment 2: In round 3, an explanation of the number of participants for different tables might help people interpret the tables correctly

In round 3, KCG included 60 people. They divided the 60 people into four conditions.

In reporting results, KCG is careful to give the "n" (number of participants) whose data contributed to each result. However, it is sometimes difficult to understand why "n" is the number it is for a specific result.

In particular, table 12 on page 14 of the round 3 report, has columns for all the conditions of round 3 (grouped as conditions 1 and 2; conditions 3 and 4). The column for conditions 1 and 2 is based on 29 people; but the column for conditions 3 and 4 is based only on 15 people – totaling 44 people for round 3. An explanation of why the other 15 people of the 60 people in round 3 are not included in this table would help.

Comment 3: In the appendix of testing materials for round 4, the title pages introducing the forms are reversed

Volume II of the 2004 report includes all the materials used in rounds 4 and 5 of the testing.

A minor glitch: The title on page 23 says: Round 4 Good Faith Estimate (GFE) Forms without Yield Spread Premium (YSP). However, the examples that follow are, in fact, the forms *with* YSP. Then, the title on page 37 says: Round 4 Good Faith Estimate (GFE) Forms with Yield Spread Premium (YSP). The examples that follow are the ones *without* YSP. Page 23 belongs at page 37, and vice versa.

Janice (Ginny) Redish, Ph.D.

Dr. Janice (Ginny) Redish is President of Redish & Associates, Inc., a small, woman-owned consultancy in Bethesda, Maryland. For more than 30 years, Dr. Redish has been helping clients develop and evaluate documents that meet both the client's business goals and the needs of the people who use the document.

Dr. Redish is a graduate of Bryn Mawr College and holds a Ph.D. in Linguistics from Harvard University.

From 1977 to 1992, Dr. Redish worked full-time at the American Institutes for Research (AIR) in Washington, DC. From 1978 to 1981, she directed the $1.5 million, federally-funded Document Design Project, bringing plain language to government forms, regulations, and other documents. In 1979, she founded AIR's Document Design Center where she and her colleagues studied the problems that people have with public documents and helped government agencies and private clients develop model documents. She directed the Center through the 1980's, also serving as Institute Director and as a Vice President of AIR.

Research with users was part of the process that Dr. Redish introduced from the beginning of her work at AIR. That research included field studies to understand how people use documents and usability testing of draft documents. For her work in this area, Dr. Redish has been called the "mother of usability" and has been featured as one of the "pioneers of usability." See www.upassoc.org/upa_publications/upa_voice/volumes/2005/june/ginny_redish.html and http://www.adlininc.com/uxpioneers/ginny_redish.html.

Dr. Redish is co-author of *A Practical Guide to Usability Testing* (with Joseph Dumas, Intellect Ltd., first edition, 1993; revised edition, 1999). The book remains a classic in the field; it is still in print and continues to be used to train new professionals in the field.

Dr. Redish's other books include:

- *Guidelines for Document Designers* (with AIR colleagues, 1981)

- *User and Task Analysis for Interface Design* (with JoAnn Hackos, John Wiley & Sons, 1998)

- *Letting Go of the Words – Writing Web Content that Works* (Morgan Kaufmann / Elsevier, 2007)

In addition, Dr. Redish serves on the editorial board of four journals and has published numerous papers and book chapters on various aspects of usability, task analysis, accessibility, document design, plain language, and writing for the web.

Dr. Redish's government clients have included the Department of Health and Human Services, the Department of Transportation, the General Services Administration, as well as agencies within federal departments, including the Federal Aviation Administration, the Food and Drug Administration, the National Cancer Institute, other institutes within the National Institutes of Health, and more than a dozen agencies in Washington state. Her non-government work has been with AARP, American Airlines, Hermes Software (Slovenia), Hewlett-Packard, Hughes Network Systems, Marriott International, Nokia (Finland), SAP (Germany), and Xerox, among many others.

Her work in plain language and usability has brought Dr. Redish many awards. She is also a Fellow of the Society for Technical Communication and a past member of the Board of Directors of the Center for Plain Language, the Society for Technical Communication, and the Usability Professionals' Association.

(For more details, go to www.redish.net.)