

After the Audit: Improving Residential Energy Efficiency Assessment Reports

Reuven Sussman, Maxine Chikumbo, and Natalia Miller
April 2019
Report B1901

© American Council for an Energy-Efficient Economy
529 14th Street NW, Suite 600, Washington, DC 20045
Phone: (202) 507-4000 • Twitter: @ACEEEDC
Facebook.com/myACEEE • aceee.org

Contents

| | |
|---|-----|
| About the Authors..... | iii |
| Acknowledgments..... | iii |
| Executive Summary | iv |
| Introduction..... | 1 |
| Purpose and Types of Home Energy Assessments..... | 2 |
| Creating and Presenting the Assessment Report | 3 |
| Previous Social Science Research on Home Energy Assessments | 4 |
| Previous Research on Creating Assessment Reports..... | 5 |
| Project Overview and Goals..... | 6 |
| Research Questions..... | 6 |
| Content Analysis | 7 |
| Elements of the Reports and What They Mean | 7 |
| Readability of Sample Reports | 13 |
| Expert Reviews | 14 |
| Expert Ratings of Assessment Reports | 14 |
| Experts' Comments on Strengths and Weaknesses of Reports | 16 |
| Eye-Tracking Study..... | 17 |
| What Homeowners Looked At in the Reports..... | 18 |
| Survey of Eye-Tracking Study Participants | 21 |
| Customer Survey | 23 |
| Satisfaction with the Assessment..... | 23 |
| The Report..... | 23 |
| Acting on the Report..... | 24 |
| Customers' Actions After Receiving Assessments | 25 |

| | |
|--|----|
| Customer Input for Creating an Ideal Report..... | 26 |
| Takeaways from Customer Survey | 28 |
| Overall Findings and Recommendations..... | 28 |
| What Are the Most Common Elements Used in Residential Energy Efficiency Assessment Reports, and What Are the Most Common Recommendations? | 28 |
| What Are the Most Attention-Catching Elements of the Reports? | 29 |
| Which Elements of the Reports Are Most Persuasive? | 29 |
| What Is the Preferred Length and Level of Detail of a Residential Energy Efficiency Assessment Report? | 30 |
| Which Metrics Do Recipients Most Prefer for Describing Costs and Benefits/Savings in the Reports? | 30 |
| How Can the Process of Presenting the Report Be Improved? | 31 |
| How Do Homeowners Use the Reports? | 31 |
| Discussion..... | 31 |
| Limitations and Future Directions..... | 32 |
| Conclusions..... | 32 |
| References..... | 33 |
| Appendix A. Methodology | 37 |
| Content Analysis..... | 37 |
| Expert Review..... | 38 |
| Eye-Tracking Study | 39 |
| Customer Survey..... | 42 |
| Appendix B. Qualitative Analysis of Individual Expert Reviews..... | 44 |
| Appendix C. Eye-Tracking Results for Home Energy Assessment Reports (Five-Page Abridged Versions) | 45 |
| Appendix D. Mediation Analysis of Verbal Presentation, Satisfaction with Form of Presentation, and Likelihood of Action..... | 48 |

About the Authors

Reuven Sussman conducts research on energy efficiency behavior-change programs and organizes the annual conference on Behavior, Energy & Climate Change (BECC). He is the author of academic papers and book chapters on the psychology of climate change, behavioral interventions to encourage energy efficiency, and the psychological determinants of pro-environmental behavior. Reuven earned a doctor of science in social and environmental psychology from the University of Victoria (Canada).

Maxine Chikumbo leads ACEEE's efforts in communications analytics, social media, and marketing and manages digital marketing activities for the Behavior, Energy & Climate Change (BECC) conference. She also conducts research and has coauthored several reports on the behavior and human dimensions of energy use. Maxine holds a BFA from the Cooper Union for the Advancement of Science and Art.

Natalia Miller is pursuing a bachelor's degree in chemical and biological engineering at Princeton University. She interned with the ACEEE Behavior and Human Dimensions program in 2018.

Acknowledgments

The authors gratefully acknowledge the sponsors, experts, external reviewers, internal reviewers, and colleagues who supported this report. This study was made possible through the generous support of Commonwealth Edison, New York State Energy Research and Development Authority (NYSERDA), National Grid, Independent Electricity System Operator (IESO), US Department of Energy, and US Environmental Protection Agency. Eye-tracking services were provided by EyeSee, a company specializing in online eye-tracking research using proprietary software. Special thanks go out to all of the home energy assessors, behavioral scientists, and marketing/graphics specialists who served as reviewers on our expert panel. Our external reviewers and panelists included Bill Morgan, Jeff Flaherty, Rob Minnick, Ted Kidd, Gary Boyer, Greg Thomas, Rhett Major, Casey Murphy, Mary Steffel, Ezra Markowitz, Mithra Moezzi, Daniel Benkendorf, Jennifer Tabanico, Amanda Carrico, Kim Wolske, Rick Larrick, Harrison Grubbs, Peter Troust, Rebecca Rabinson, Marcela Gara, DeeAnn Budney, Amanda Dwelley, Beth Karlin, Nic Crowder, Michelle Ackmann, Bill Keating, Marsha Walton, Romilee Emerick, Joe Bryson, Robert Merrin, and Cyrus Shaoul. External review and support do not imply affiliation or endorsement. Internal reviewers included Steven Nadel, Naomi Baum, and Jennifer Amann. We also gratefully acknowledge the assistance of Michel Wedel from the University of Maryland. Last, we would like to thank Fred Grossberg for developmental editing and managing the editing process; Elise Marton, Sean O'Brien, and Roxanna Usher for copy editing; Eric Schwass for graphics design; and Casey Skeens, Kate Doughty, and Wendy Koch for their help in launching this report.

Executive Summary

KEY TAKEAWAYS

- Reports should present all key information in the first four to five pages and provide more detailed information (of interest to highly engaged homeowners or their contractors) in supplementary sections or an appendix.
- Persuasive reports compare the current home to average and potential future homes (after improvements), use vivid language, show a variety of benefits (e.g., health, comfort, and cost savings), explain clear next steps, and provide all the information a homeowner needs to act on recommendations.
- Assessors should prioritize their recommendations on the basis of customers' concerns and needs, add explanations for their recommendations, and include a personalized narrative summary (a first-person description of results and recommendations, written in the assessor's own voice).
- Recipients of residential efficiency assessment reports are more likely to follow through on recommendations if assessors make them accessible and personally engage with recipients (e.g., by walking them through the house and going over the written report with them).
- Assessors should write their reports as clearly and simply as possible, removing unnecessary elements and using language that is easy to understand, preferably at about an eighth-grade reading level.
- Reports should feature good design, including a consistent structure, a mix of text and visual elements, several high-quality photos (featuring good content, composition, and lighting), explanations of the photos, and tables and graphs that simplify information.

RESIDENTIAL EFFICIENCY ASSESSMENTS

Energy efficiency assessments are a key piece of most utilities' residential demand-side management portfolios. Almost nine million US households had been assessed as of 2015.¹ The way in which residential energy efficiency assessors present results can influence homeowner satisfaction with assessments and the likelihood that they will take action. The overarching goals of this project are to describe current practices in assessment report writing, characterize the responses of homeowners to receiving reports, and identify opportunities for improving them.

¹ Energy Information Administration, "Table HC2.3: Structural and geographic characteristics of U.S. homes by year of construction, 2015," *Residential Energy Consumption Survey (RECS)*, February 2017. www.eia.gov/consumption/residential/data/2015/hc/php/hc2.3.php.

METHODOLOGY

This project is the most thorough analysis of assessment reports conducted to date. We used a mixed-method approach, blending qualitative and quantitative methods, to triangulate our results and ensure reliability. We took the following steps:

- *Content analysis*: Report review, element coding, and readability calculation
- *Expert review*: Evaluation of reports by content specialists
- *Eye-tracking study*: Observation of homeowner eye movements while reading
- *Customer survey*: Survey of customers receiving residential assessments

CONTENT ANALYSIS

The research team analyzed 45 reports, coding 65 elements of each and calculating their readability scores. The most common elements of the reports were text, photos, diagrams, and tables and graphs. The sample reports we collected were somewhat long (averaging 14 pages), with dense text (224 words per page) and high language difficulty (11th-grade level, rather than the 8th-grade level that we recommend). On average, reports had 1.25 images per page (photos of home, stock photos, thermal images, or diagrams), of which 56% were tailored to the assessed home. We found that 73% of reports used tables to present information, and 29% included narrative summaries. These elements are particularly effective for conveying recommendations.

For measuring benefits, the most common metrics included in the reports were annual savings (60%), up-front costs (58%), and rebate or discount information (42%). Less common metrics included health inspection results (40%), payback period (31%), actual energy saved (27%), and carbon footprint reduction (20%).

EXPERT REVIEWS

Teams of 3 content experts (behavioral scientists, efficiency assessors, and marketing and design specialists) evaluated the 42 highest-scoring reports (based on ACEEE preliminary ratings). Reviewers generally agreed that they could be greatly improved. In particular, assessors could strengthen the recommendations sections by making them easy to find, presenting them clearly and simply, improving the content (e.g., by avoiding unnecessary recommendations), prioritizing suggestions, and explaining this prioritization. The experts also noted that assessors should tailor reports to the individual homeowner, remove barriers to follow-through (e.g., by providing specific information on what the homeowner should do next, along with relevant resources), and avoid technical language and unnecessary information. They said that narrative summaries of results and recommendations, complemented by tables summarizing this information are particularly useful and important for homeowners; these should be placed within the first four to five pages of the report, including multiple metrics of benefits and explanations.

Experts gave higher scores to reports that included payback periods, but they were split as to the persuasiveness of this information. Some found the metric useful for assessing benefits, but others suggested that the calculations were inaccurate or that they focused too much attention on the financial aspects of homeowners' decisions. In general, experts

lamented the heavy emphasis on financial motivation and recommended presenting multiple benefits instead. On average, reports presented three benefits metrics each.

EYE-TRACKING STUDY

We condensed three of the top-rated reports to five pages each and included them in an eye-tracking study. EyeSee administered the study to a nationally representative sample of homeowners. We found that homeowners paid attention to critical information in the reports, especially if it came earlier rather than later and was placed at the top of a page. Although they took in most text and visual elements, they particularly attended to summary tables and narrative summaries. In general, attention was dictated more by the structure of the page than by how interesting or important the information was. Including a variety of elements is effective, but having too many can result in homeowners skipping some information. Based on our limited sample, we suggest including no more than four to five elements per page. However this would vary based on size, complexity, and spacing of elements.

CUSTOMER SURVEY

We surveyed customers in New York and Illinois who received assessments to learn what actions they took after their assessments and how the report process could be improved. Survey findings suggested that assessors should present assessment results both in writing and in person, which corroborated what assessors on our expert panel told us. Customers who received an oral presentation were more satisfied with how the results were conveyed than were those who received only written information. They were also more likely to act.

Homeowners were generally satisfied with their reports (rating them 82 out of 100) but a little less satisfied with the overall assessment (78 out of 100). They reported upgrading for reasons of comfort, energy savings, or cost savings. Reasons given for *not* upgrading were financial (upgrades were too expensive, or homeowners were not sure about costs and savings). Most often, customers saved the reports for their records (41%) and/or installed upgrades (34%), usually with the help of a contractor (they were twice as likely to hire a contractor than to do it themselves). Some 27% of participants said they applied for a rebate, and 14% said they contacted a contractor for more information. The most common upgrade was installing insulation (51%), followed by weather-stripping and air sealing (35%).

RECOMMENDATIONS

We suggest the following to maximize the persuasiveness of efficiency assessment reports and encourage homeowners to follow through on recommendations.

- Present the report in person as well as in written form.
- Use simple, clear language and personalized information instead of boilerplate in the written report.
- Maintain a simple and consistent structure.
- Provide key information in the first four to five pages, followed by more details in supplementary material at the end.
- Prioritize recommendations, with explanations, based on customer concerns and needs.

- Summarize results and recommendations in tables and in personalized narrative summaries.
- Present a variety of upgrade benefits including health, comfort, and cost savings.
- Highlight social norms or offer comparative information, for example the customer's home energy use versus that of the average home and potential future homes (with improvements).
- Use vivid, metaphorical language to bring the report to life.
- Include high-quality images (with good content, composition, lighting, and so on) and explain them with notations.

Introduction

Residential electricity use accounts for the largest proportion (37.4%) of American electricity consumption (EIA 2018a). It may also represent one of the biggest opportunities for energy savings. National Renewable Energy Laboratory estimates the US residential building stock could save \$49 billion in annual energy costs through efficiency improvements (NREL 2017).

Building new homes that are more energy efficient is only part of the solution; upgrading existing homes is also critical. Old homes make up a large proportion of the American housing stock and tend to be relatively inefficient. A Swiss study notes that buildings constructed between 1947 and 1979 generally consume more energy than those built earlier or later (Aksoezen et al. 2015). In the United States, 36% of buildings were erected around this time (EIA 2018b). These buildings might be prime targets for energy assessments and energy efficiency upgrade work.

Changes in household behaviors can impact energy consumption in existing homes, and a one-time investment in energy upgrades may be the most impactful behavior of all (Gardner and Stern 2008). Many utilities use home energy assessments as a strategy to encourage installation of energy-saving upgrades and reduce energy demand. In 2013, a large meta-analysis of experimental studies to reduce energy use found that interventions such as home energy assessments were more effective than other popular interventions including peer comparison feedback (Delmas, Fischlein, and Asensio 2013). As of 2015, 8.9 million households in the United States had received one of these assessments, leading to average energy reductions of 5% or more (EIA 2018b; Delmas, Fischlein, and Asensio 2013; Alberini and Towe 2015).

Nevertheless, home energy assessments are not as effective as they could be because conversion rates (the proportion of customers who complete recommended upgrades) are sometimes low (e.g., Breukers et al. 2009; Jaffe and Stavins 1994). Frequently participants go through the assessment process but either do not complete the recommended upgrades or complete only the cheapest ones with the shortest payback periods (e.g., Palmer, Walls, and O’Keeffe 2015). Conversely, many homeowners complete upgrades without an assessment, which calls into question the necessity of assessments for promoting upgrades (Frondel and Vance 2013).

In a 2013 study, 71% of assessors estimated that American homeowners installed at least one recommended upgrade “fairly often” or “always,” but only 1% of homeowners followed through on all recommendations (Palmer et al. 2013). Surveys of assessment recipients find similar results. One home energy assessment program, the E-Conservation Residential Audit Program, achieved mean annual savings of \$357 per customer, with 84% implementing at least one low- or no-cost change and 55% implementing other upgrades (Kirby et al. 2014). A program in Pittsburgh led by a student group, the National Energy Leadership Corps, found that 85% of all customers adopted at least one recommendation, but overall only 30% of recommendations were adopted (Ketchman et al. 2016). Other studies have found similar adoption rates of 30–85% (Fuller et al. 2010; Ingle, Moezzi, and Lutzenhiser 2012; Palmer et al. 2013; Murphy 2014). All of this indicates that home energy assessments are somewhat effective but could be improved, particularly for encouraging the adoption of higher-cost recommendations.

Homeowners most commonly cite financial factors as reasons for upgrading or not upgrading their homes (Sussman and Chikumbo 2017). Still, nonfinancial factors can also be important. Customers' personal circumstances, values, attitudes, and interests can motivate (or inhibit) action. Programs offering identical incentives can differ tenfold in effectiveness, based on the strength of their marketing campaigns (Stern et al. 1986).

PURPOSE AND TYPES OF HOME ENERGY ASSESSMENTS

A home assessment is the main element of a home energy audit. The primary purpose of a home assessment is to evaluate the current state of a residence's energy efficiencies and identify opportunities for efficiency improvements. Homeowners typically conduct assessments either with the assistance of a residential energy efficiency expert or a contractor with specific expertise (e.g., HVAC or insulation), or alone, usually with the help of a computer-based tool or website) (Riley et al. 2012, cited by Sprehn et al. 2015). In some cases, the professional assisting with the assessment also may be qualified to install the recommended measures and, as such, may be particularly motivated to encourage homeowners to act. In these situations, honest and unbiased assessments are important for gaining credibility and building trust from homeowners, ultimately helping to influence homeowners' investments in energy efficiency upgrades (Sussman and Chikumbo 2017).

Homeowners can conduct several types of assessments. One taxonomy includes four types: "self" (with the help of an online or computer-based calculator), "assisted" (with the help of a specialist in one type of installation or a general nonexpert such as a member of a community group), "traditional" (guided by an energy efficiency expert), and "deep" (guided by an expert looking for savings of 50% or more) (Riley et al. 2012, cited by Sprehn et al. 2015). In our discussions with assessors, we also noted two potential additional types of assessment: those that are conducted only for the direct installation of free or subsidized upgrades, and those conducted purely for energy score calculations.

Although many assessment programs include a direct installation component, some programs are entirely designed around the installation of specific measures. Direct installation assessments occur when utilities offer free upgrades; more extensive free measures to homeowners meeting certain criteria, such as low-income residents qualifying for weatherization programs (DOE 2019b); or low-cost measures to all customers. In those cases, home energy professionals may go to customers' homes, conduct assessments, and then directly install qualified upgrades. Assessors typically do not recommend additional measures. Reports for these assessments may be cursory and lack detail. Assessments (sometimes self-assessments) may also be conducted to determine a home energy score or similar metric (DOE 2019a). Again, homeowners may receive a report, but it will typically highlight only the energy score. Because of their cursory nature, these types of assessments and reports are not the focus of this paper.

Most assessments, and especially those that are more rigorous and detailed, tend to proceed through four major steps (Sprehn et al. 2015):

- Gathering homeowner facts
- Gathering facts about the residence
- Identifying and analyzing energy efficiency opportunities
- Presenting results to homeowners

Assessors learn these steps through certification programs and organizations such as the Building Performance Institute, Home Performance with ENERGY STAR®, and the Residential Energy Services Network.

CREATING AND PRESENTING THE ASSESSMENT REPORT

The final step of the assessment procedure, presenting results, usually includes some form of written report. Although our focus in this paper is the structure, appearance, and writing of the report, we learned from experts and homeowners that how the assessor presents the report also matters. When assessors personally go over the report with homeowners, they are more likely to influence follow-through. One assessor put it bluntly: “Any report delivered without a presentation is toilet paper.” This meshes with previous research showing that guiding customers through information barriers, decision-making barriers, and transactional barriers increases assessors’ success (Billingsley, Stratton, and Fadrhonc 2016).

To the best of our knowledge, there are few written guides describing best practices for creating residential efficiency assessment reports from a sales and marketing perspective. *Selling Energy* (2014) by Mark Jewell is a comprehensive guide to sales and marketing for energy efficiency. Nevertheless, the book does not focus specifically on the design and creation of assessment reports. The lack of broad knowledge on this topic leads assessment reports to vary greatly in terms of length, structure, language, and use of graphics and images. Home energy assessors are good at diagnosing and addressing technical opportunities for efficiency but are generally less experienced in effectively presenting their recommendations from a marketing standpoint (e.g., Gonzales, Aronson, and Constanzo 1988). This is not surprising, given that home energy assessors are hired for their energy expertise as opposed to their communication or design skills.

Report Generation Software

Fortunately, assessors can purchase software that allows them to input their findings and automatically create well-designed reports. These can facilitate the process of creating effective materials without the need to learn and understand design principles; they can also reduce the time required to generate reports after an assessment. Examples of these programs include Energy Optimizer, Snugg Pro, and TREAT. The advent of a common data standard, HPXML, is now further facilitating automatic report generation.² This new standard allows a home energy assessor to collect data with a variety of tools and use

² For more information, see www.hpxmlonline.com/.

different options of software programs to export the data into a report. Some organizations, such as the New York State Energy Research and Development Authority (NYSERDA), require their assessors to use approved software programs (those capable of generating HPXML data) and to upload their assessment results in HPXML form to their database.

Automatic report generation software is an excellent step forward, allowing assessors to focus on assessing rather than designing reports. But it is not the complete solution, because of two main drawbacks. The first drawback is that the software can be expensive. Currently, software solutions usually charge per assessment report generated, sometimes with discounts given for higher volumes. This system therefore favors larger companies that process greater numbers of assessments. Small companies or individual contractors may therefore prefer to continue creating their own reports, especially if they are unlikely or unable to earn additional income by doing the home upgrades themselves.

The second drawback of report generation software is that assessments still need human input. For example, as we discuss later in this report, personalized narrative summaries, as well as thermal images and custom-tailored results, are key elements that make reports effective. This means that, as good as the software might be, the author of the report must still work to make the report conform to best practices. Assessors should exclude sections that are unnecessary, reorder sections to emphasize solutions to homeowner concerns, use personal and nontechnical language, write a narrative summary, and acquire thermal images to put into the report. Software, in itself, does not do all of these functions.

PREVIOUS SOCIAL SCIENCE RESEARCH ON HOME ENERGY ASSESSMENTS

Program implementers can increase the likelihood that homeowners will upgrade their homes either by increasing the likelihood that homeowners will have an assessment, or by increasing the effectiveness of the assessment and the presentation of results or recommendations.

Two studies have tested theory-based approaches for encouraging homeowners to participate in assessments. One study found that sending assessment invitations to homeowners with social comparison information makes them more likely to request one than those who receive invitations without the comparison (Holladay et al. 2016). The other found that when contractors send notecard reminders about assessments that use social science-derived behavioral strategies (citing social norms, emphasizing key elements, using a personal touch), the recipients are more likely to keep appointments than when the notecards do not use those strategies (Gillingham and Tsvetanov 2018).

In regard to increasing the effectiveness of the assessments themselves, Sussman and Chikumbo (2017) summarized how psychology and behavioral science can inform multiple aspects of the assessment. These include the cost of the assessment and recommended measures, direct installation of upgrades, interpersonal style and personal characteristics of the assessor, performance of specific diagnostic tests, and the method of collecting and presenting data. They also evaluated the effectiveness of highlighting various benefits of upgrading (e.g., financial, health, comfort, home value) in an online experiment that tested six message-framing strategies that assessors could use when discussing upgrade options. The outcome of the Sussman and Chikumbo 2017 research was a series of recommendations

for how assessors could talk about home energy upgrades with homeowners. It emphasized the use of behavioral insights, such as message-framing strategies that change how customers perceive costs.

This paper follows up on Sussman and Chikumbo's 2017 study. We focus specifically on how assessors can improve their reports to maximize the likelihood that customers will follow through on their recommendations. The assessment report is a key deliverable for the homeowner, one that can influence satisfaction (or dissatisfaction) with the assessment and upgrade decisions. This paper examines the assessment report as a whole, identifying common practices and how they can be changed to better capture homeowners' attention, convey information persuasively, and effect behavior change.

PREVIOUS RESEARCH ON CREATING ASSESSMENT REPORTS

Two studies have previously made recommendations for improving residential efficiency assessment reports. One of these was the above-mentioned report by Sussman and Chikumbo (2017) on message-framing strategies for home energy upgrades. The series of experiments described in that study helped determine that, in addition to mentioning financial benefit, messages emphasizing greater comfort and health could increase the intention to upgrade. Similarly, listing "no-brainer" upgrade recommendations separately from other recommendations can slightly increase homeowners' intentions to purchase midrange upgrades that are more expensive than the no-brainer recommendations. Presenting payback in terms of a specific month and year in the future could be a little more effective than talking about the number of years and months from today. And presenting the total cost of upgrades as an add-on to needed repairs, as opposed to a stand-alone investment, could increase homeowners' intention to follow through.

Regarding the assessment report, the takeaway recommendations of Sussman and Chikumbo (2017) were to tailor the report to homeowner values and concerns, to put a greater focus on nonfinancial benefits (especially if those might be of particular interest to the homeowner), and to apply the subtle message-framing strategies described above. The study provided some specific recommendations for presenting certain types of results and described overall strategies for interacting with homeowners, but it did not examine the report as a whole, nor did it investigate aspects such as overall language, look, and length of the report.

The second study was a content analysis of assessment reports (Sprehn et al. 2015). The researchers analyzed 10 reports to identify common content elements and determine which of these should be considered essential. They identified 10 sample assessment reports using a standard Internet search and coded those reports to describe 30 elements that were most common among them. Then, in a survey of 500 assessors, they asked whether each report element was "essential," "useful but not essential," "unessential," or "other," using a method developed by Lawshe (1975) to calculate the "essentialness" of different factors. This showed that four elements were essential: (1) a summary of the audit procedure, (2) a summary of recommendations, (3) building photographs, and (4) numbers (e.g., kilowatt-hours, dollars, or therms). Through their survey responses, assessors did not deem any of the other 26 elements, such as resources for further education or tax incentive programs, as "essential," "useful but not essential," or "not essential" to a statistically significant degree.

Sprehn et al. (2015) provides important insights but has several limitations. The study was conducted to identify essential sections of reports, primarily from a technical and information-provision perspective rather than a customer-centric or behavior-change perspective. Thus, its findings were limited by exclusive use of expert assessors as opposed to behavioral scientists, marketing professionals, graphic designers, or utility customers (particularly those who actually received assessment reports). Moreover, its selection of reports was limited to just 10 publicly available documents. Although most structural elements were represented, the authors could not generalize to aspects beyond those presented in the 10 reports. Furthermore, they may not have used the best reports available; one was examined by the panel of expert reviewers we convened for this study and was ranked 39th out of 42 reports. Overall, Sprehn et al. (2015) was helpful for determining content but had limited application in terms of persuasiveness and behavior change.

Project Overview and Goals

Both the Sussman and Chikumbo (2017) and Sprehn et al. (2015) studies provide useful guidance for assessors but leave several aspects of report development unexplored. The current study builds on the findings of these previous works by examining a larger sample of reports with a more diverse group of experts and homeowners and using a larger variety of analysis approaches. It seeks to go beyond asking what sections are important or what message-framing strategies should be employed, asking how reports should be designed and written overall, as well as how customers respond to receiving them.

The overarching goals of this project are to identify current practices in assessment report writing, responses of homeowners to receiving reports, and opportunities for improving reports. To do this, we took the following steps (a more detailed description of our methodology is provided in Appendix A):

- *Content analysis.* Report review, element coding, and readability calculation
- *Expert review.* Evaluation of reports by content specialists
- *Eye-tracking study.* Observation of homeowner eye movements while reading
- *Customer survey.* Survey of customers receiving residential assessments

RESEARCH QUESTIONS

Through our deep dive on residential efficiency assessment reports, we sought to answer the following questions:

- What are the most common elements used in residential energy efficiency assessment reports?
- What are the most attention-getting elements of the reports?
- What elements of the reports are most persuasive?
- What is the optimal length and level of detail of a residential energy efficiency assessment report to encourage customer follow-through?
- Which metrics do recipients most prefer for describing costs and benefits or savings within the reports?
- How can the process of presenting the report and following up be improved?

Before answering these questions, we describe the findings of each section of our research.

Content Analysis

The first part of our study involved an in-depth examination of sample residential efficiency assessment reports. As described in Appendix A, we collected 68 reports, which we narrowed to a group of 45 for coding and analysis. Nine of the reports came from the US Northeast, seven from the Midwest, 15 from the West, and eight from the South. In addition, two reports came from outside the United States, and four came from unspecified locations. Several reports were created with known software templates: four by Snugg Pro, three by Energy Optimizer, two by TREAT, and one by REM/Rate. For content analysis, we coded 65 elements of the reports, including basic characteristics, readability, and design, as well as a few aspects of specific content (e.g., presence of a health and safety assessment).

ELEMENTS OF THE REPORTS AND WHAT THEY MEAN

Given that Sprehn et al. (2015) had already conducted a content analysis on specific sections, we focused our analysis on elements that made reports persuasive, clear, and eye-catching. The tables below describe the design and content elements found in our collection of sample reports. Below, we summarize these characteristics and put them into context, given the comments from expert reviewers (described later) and eye-tracking research (also described later).

Text and Length

Reports averaged 14 pages in length and 224 words per page. Experts tended to note that most reports were made too long by unnecessary information. The number of words per page is slightly less than what would be expected on a full page of double-spaced, 12-point text (approximately 250 words). Given that each assessment report page had an average of 1.25 images, the text appeared to be somewhat dense. Twelve reports included appendixes, with an average length of three pages. Experts liked reports with appendixes or those that contained critical summary information up front followed by in-depth reference material toward the end. Effective reports had the most critical actionable information within the first four to five pages. Table 1 and figure 1 present these characteristics.

Table 1. Report text and length

| Characteristic | Average |
|-------------------------------------|---------|
| Pages per report (single-sided) | 14 |
| Pages per appendix (when included)* | 3 |
| Words per report | 2,786 |
| Words per page | 224 |

*Twelve reports included appendixes.

| Table of Contents | | |
|-------------------|---|-----------|
| 1 | Introduction | 3 |
| 1.1 | The Energy Audit | 3 |
| 1.2 | Improvement Opportunities for Your Home | 3 |
| 1.3 | Benchmarking Your Home | 4 |
| 1.4 | HERS Score | 5 |
| 2 | Your Home in its Current Condition | 6 |
| 2.1 | Health and Safety Considerations | 6 |
| 2.2 | Insulation Levels and U.S. Department of Energy Recommendations | 7 |
| 2.3 | Air Leakage Test | 8 |
| 2.3 | 2013 Federal Tax Credits | 18 |
| 2.4 | Solar Financing & Incentives | 19 |
| 6 | Appendices | 20 |
| 6.1 | Building Science – The House as a System | 20 |
| 6.2 | Lights and Appliances | 21 |
| 6.3 | Basic Energy Saving Tips | 22 |
| 6.4 | Water Efficiency Tips | 23 |
| 6.5 | Resources | 24 |

Figure 1. Table of contents for a well-structured report by Envinity. It includes key information in the first five pages and appendixes with additional information at the end.

Images

Reports had an average of 1.25 images per page (17 images in total). These included interior and exterior photos of the home, stock photos of other homes (or people working or living in them), and diagrams.³ A large majority of reports (82%) devoted less than a quarter of their space to images. The use of images can help draw and keep attention and stimulate visual interest, but our eye-tracking analyses showed that too many graphic elements could also be distracting. Graphic designers in our panel noted that photos need explanation (e.g., arrows or circles pointing out important details) and that the quality of photos – in terms of content, composition, lighting, and so on – can strongly impact the overall perception of quality and credibility of the assessment. All expert groups also noted that personalization was critical for the reports to be effective. Approximately 56% of all images were personally tailored to the customers' homes. Table 2 and figure 2 present these characteristics.

Table 2. Images

| Characteristic | Average/ frequency |
|--|-----------------------|
| Images per report | 17 |
| Images per page | 1.25 |
| Percentage of all images that were personally tailored to customers' homes | 56% |
| Reports that included thermal images* | 38% |

³ We considered a group of images presented together to be one large image. Graphs, charts, and tables were not considered images for this calculation.

| Characteristic | Average/ frequency |
|---|-----------------------|
| Reports that devoted less than ¼ of space to images | 82% |
| Reports that devoted ¼ to ½ of space to images | 11% |
| Reports that devoted ½ to ¾ of space to images | 4% |
| Reports that devoted more than ¾ of space to images | 2% |
| Reports that did not use color images | 7% |
| Reports that did not include any images | 2% |

* 99% of thermal images were tailored to customers' homes.



Figure 2. A good use of images, with clear photographs of the customer's home (rather than stock photos) alongside an explanation of the energy issue being depicted. *Source:* Assessment report by Hassler Heating and Air Conditioning.

Tables and Graphs

Most reports included some type of table or graph. Seventy-three percent used summary tables to present information such as recommendations and home diagnostics, and 51% included a graph (usually to show electricity usage). Our eye-tracking study showed that summary tables attract and hold attention, often retaining viewers for longer than complete paragraphs even though they included fewer words. Table 3 and figure 3 present these characteristics.

Table 3. Tables and graphs

| Characteristic | Frequency |
|--|-----------|
| Reports that included a graph* | 51% |
| Reports that used summary tables to present information | 73% |
| Reports that used tables to summarize recommendations | 62% |
| Reports that used tables to summarize problems in the home | 29% |
| Reports that used tables to present rebate offers and packages | 29% |
| Reports that used tables to summarize direct installations | 16% |

* The most common type of graph was a bar graph depicting electricity usage.

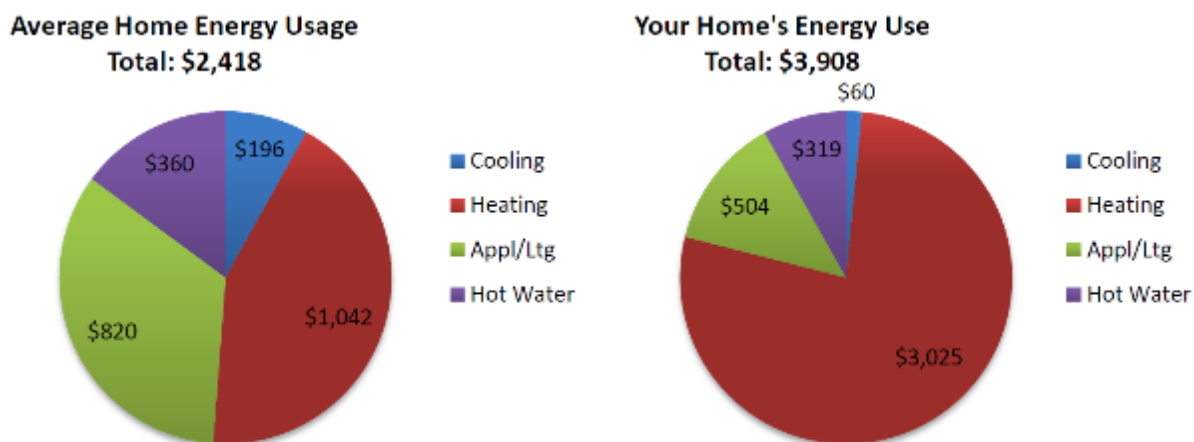


Figure 3. Effective graphs simplify and clarify information. Presenting information about the current home is more effective if it is presented in the context of other homes or the average home. This strategy also helps persuade the customer to take action by leveraging social norms. *Source:* Assessment report by Envinity.

Recommendations Section

Relatively few reports, 29%, included a narrative summary of all the findings and recommendations. Almost all of these (12 of 13) put the summary at the beginning of the report. Our panel of experts commented on how a personalized summary at the start of the report was important to increase trust, sincerity, clarity, and credibility. Our eye-tracking study corroborated this, showing that homeowners spent more time viewing and focusing on narrative summaries than on other, similar areas of text. When assessors listen to homeowners and reflect their concerns in the narrative summary using an honest tone and clear, simple language, homeowners will be more motivated to act.

Recommendations could be improved by providing behavior-change tips and clearly identifying next steps. Behavior-change tips were suggested in only 33% of the reports, and clear next steps (including all the information that customers would need to follow through) were described in only 56% of the reports. Behavioral scientists explained that removing barriers to action should be an important goal of these reports, and providing information for next steps is one way to do so. If the information for next steps is difficult to find (or absent), then the likelihood of action is reduced. Table 4 and figure 4 focus on recommendations.

Table 4. Recommendations

| Characteristic | Frequency |
|--|-----------|
| Reports that included a narrative summary of findings and recommendations* | 29% |
| Reports that presented clearly defined next steps for the customer | 56% |
| Reports that included behavior-change tips for energy savings | 33% |

*Of the reports that included a narrative summary, 92% put it at the beginning of the report.

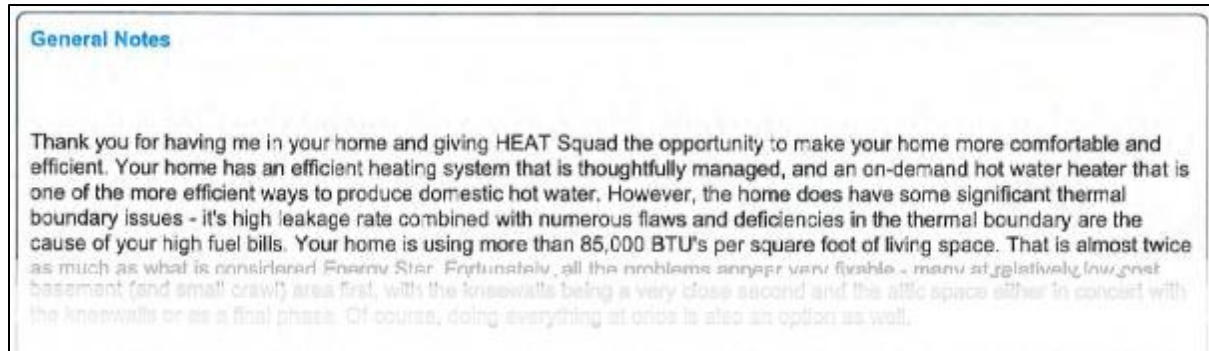


Figure 4. Personal narrative summary. *Source:* Assessment report by Neighborworks H.E.A.T. Squad.

Savings Metrics and Benefits

The most common metrics for measuring benefits that assessors included in their reports were annual savings (60%), up-front costs (58%), and rebate or discount information (42%). Other metrics, such as health inspection results (40%), carbon footprint reduction (20%), and energy saved (e.g., in kWh) (27%) were less common. In their comments, assessors observed that too few reports promoted nonfinancial benefits. Nevertheless, experts agreed that clear financial information was important to include and that generally presenting multiple metrics was helpful. On average, reports offered three different savings metrics (listed in table 5 below). Multiple metrics can help make reports more persuasive, especially if the information is presented in the context of other homes, the average home, or the customer's potential future home (after improvements). Customers can find and use the information that matters most to them and make an informed decision.

Payback period was presented in 31% of reports, mostly for items with immediate or short-term paybacks (less than four years). This was a somewhat controversial metric, as some assessors noted that these were unrealistically short and that (correctly) citing longer periods would be very unpersuasive. Behavioral scientists felt it was important to provide payback period because it simplifies financial calculations. As discussed below, a statistical analysis found that payback period was one of two factors that independently predicted higher overall expert ratings of report quality. Table 5 and figure 5 show these features.

Table 5. Savings metrics and benefits presented to customers

| Characteristic | % of reports that included metric |
|--|-----------------------------------|
| Financial | |
| Annual savings | 60% |
| Up-front costs | 58% |
| Rebate or discount information | 42% |
| Payback periods | 31% |
| Savings-to-investment ratios (SIR) | 22% |
| Lifetime or 10-year savings | 7% |
| Maintenance costs | 4% |
| Return on investment (ROI) | 4% |
| Energy | |
| kWh savings | 27% |
| Therm or natural gas savings | 8% |
| Home energy score | 7% |
| HERs index | 7% |
| Environment and water | |
| Carbon footprint | 20% |
| Water savings (gallons) | 9% |
| Health and safety | |
| Combustion safety | 49% |
| Carbon monoxide | 47% |
| Specific health and safety recommendations | 38% |
| Social comparison | |
| Comparison to other homes or “average home” | 20% |
| Benefits | |
| Cost savings | 72% |
| Comfort | 72% |
| Health | 42% |
| Home value or upgrades as an investment | 13% |
| Environmental sustainability, carbon footprint, or greenhouse gas emissions reductions | 27% |

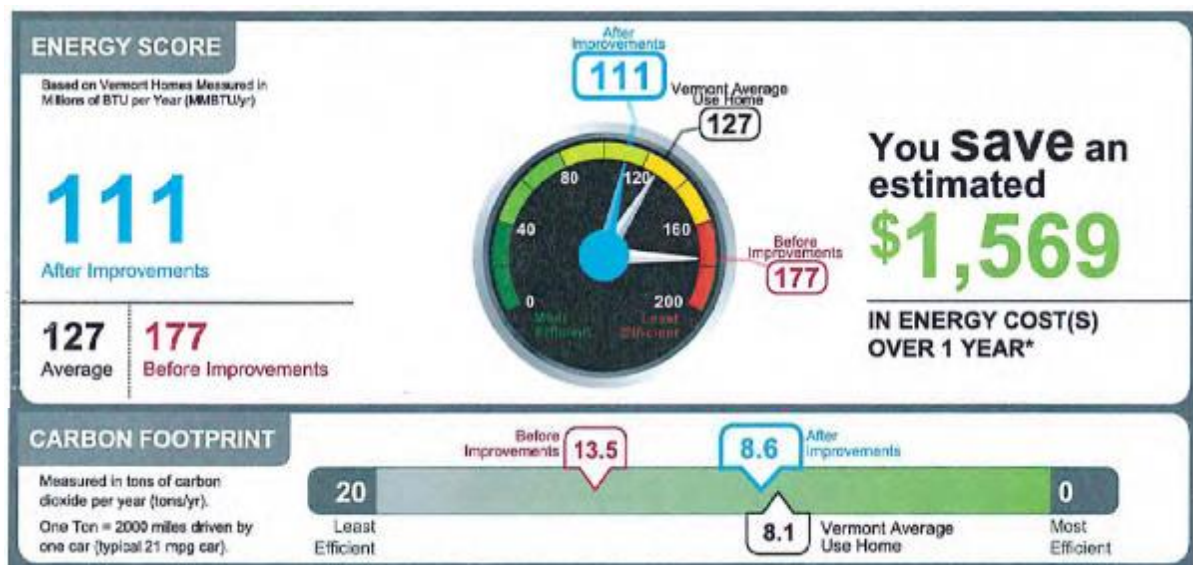


Figure 5. Presentation of multiple metrics for financial and nonfinancial benefits. *Source:* Assessment report by Neighborworks H.E.A.T. Squad.

READABILITY OF SAMPLE REPORTS

Simplifying written marketing materials (e.g., pamphlets encouraging residents to enroll in energy conservation programs) to improve their readability can make them more effective (Wong-Parodi, de Bruin, and Canfield 2013). We calculated objective readability scores for each home energy assessment report using procedures described in Appendix A. Based on the combined scores of multiple readability metrics, our sample assessment reports required, on average, an 11th-grade reading level (results ranged from 8th grade to the second year of college). This is higher than recommended and may make it difficult for assessment recipients to understand and engage with their assessments.

The readability web tool that we used, Readable.io, explains that for content to be “readable by 85% of the general public, you should aim for a readability score of Grade 8 or better” (by “better,” they mean lower) (Readable 2019). However we could not confirm this through other sources. Given that 92% of American homeowners have at least a high school education (Census Bureau 2017), an 11th-grade reading level might be considered reasonable. That said, the last time the National Assessment of Adult Literacy survey was conducted, it found that 43% of American adults had only a “basic” or “below basic” level of literacy (NCES 2003). Although this is not mapped onto the “grade level” concept of literacy, one would surmise that materials written at an 11th-grade level might be difficult for these readers to fully understand.⁴ According to Johns and Wheat (1984), American newspapers are usually written at about an eighth-grade level.⁵ The readability web tool we

⁴ These statistics come from the most recently available National Assessment of Adult Literacy in 2003. But at that time results did not change significantly from 10 years prior (nces.ed.gov/naal/kf_demographics.asp). This suggests that, even if literacy rates have changed since that time, it is unlikely that they changed dramatically.

⁵ Reading level varies according to topic, source, and grading tool (Johns and Wheat 1984).

used gave our sample of assessment reports an average letter grade between B and C (the lowest possible was D).

Expert Reviews

We recruited a panel of experts from three disciplines to review our sample of reports: eight home energy assessors, eight behavioral scientists, and six graphic designers or marketers. Each report was reviewed by one expert of each type, using standardized review forms. Each expert conducted approximately seven reviews. In open-ended questions, the forms asked reviewers to provide qualitative comments on strengths and weaknesses of each report (described in Appendix B); other questions asked for quantitative ratings of several aspects of the reports. On completion of the reviews, we convened three focus groups to discuss the reports: one for behavioral scientists, one for assessors, and one for graphic designers. The procedures for recruiting panelists, assigning reviews, coding results, and analyzing outcomes are described in Appendix A. The specific qualitative analyses from individual reviews and focus groups are available in Appendix B.

EXPERT RATINGS OF ASSESSMENT REPORTS

In their individual reviews, experts rated the assessment reports in general and on four specific dimensions, presented in table 6. Overall, reviewers from all three disciplines felt that there was significant room for improvement in the assessment reports. Experts individually rated qualities of each assessment report on a seven-point scale, with 1 being the lowest score and 7 the highest. Overall quality was rated 3.94 out of 7, understandability received a rating of 4.16, persuasiveness was rated 3.44, interestingness and appeal received a rating of 3.86, and the summary and recommendations sections received a rating of 3.92. These varied by type of reviewer, with behavioral scientists usually awarding slightly higher average scores than assessors or graphic designers.

Table 6. Experts' mean scores for assessment reports

| Raters | Overall | Understandability | Persuasiveness | Interestingness and appeal | Summary and recommendation sections |
|-----------------------|---------|-------------------|----------------|----------------------------|-------------------------------------|
| Graphic designers | 3.86 | 3.95 | 3.08 | 3.67 | 3.93 |
| Assessors | 3.85 | 4.3 | 3.53 | 3.85 | 3.7 |
| Behavioral scientists | 4.17 | 4.29 | 3.68 | 4.07 | 4.29 |
| Combined average | 3.94 | 4.16 | 3.44 | 3.86 | 3.92 |

Overall report quality was correlated with understandability ($r = .82, p < .001$), persuasiveness ($r = .89, p < .001$), quality of summary and recommendations ($r = .89, p < .001$), and interestingness and appeal ($r = .89, p < .001$). However these were closely related to one another. To determine the best unique predictor of overall quality, we conducted a regression analysis, $F(4, 37) = 164.63, p < .001, R^2 = .95$. Interestingness and appeal ($\beta = .45$), as well as quality of summary and recommendations ($\beta = .38$), accounted for the most

unique variance in the rating of overall quality of the reports, suggesting that these two aspects may be particularly important.

Social Comparison and Payback Period

We conducted exploratory analyses to determine if any specific elements that we coded in part one of the project predicted experts' ratings of overall quality. Generally, we found the effect to be diffused across elements and that the combination of elements was likely important. However two specific elements, payback period and social comparison information, also significantly predicted overall ratings on their own.

We found that reports that mentioned payback period were rated higher ($M = 4.44$) than those that did not ($M = 3.69$), $t(40) = 2.25$, $p = .03$. In their comments, some reviewers noted that the metric was helpful, but others said payback period in some of those reports was a weakness (commenting, for instance, that unrealistically short periods were probably unpersuasive, or that the period was too long to be compelling). Generally, behavioral scientists were more likely to think payback period was helpful (because it simplifies financial calculations for homeowners), whereas assessors found it unhelpful (because the periods were either quite long or not calculated correctly). Sussman and Chikumbo (2017) suggest that when payback period is long, assessors should consider de-emphasizing the statistic and focusing on nonenergy benefits, which are not usually accounted for in financial payback calculations. Homeowners invest in upgrades for a variety of reasons beyond financial savings, and they often welcome discussions of nonfinancial benefits. They may choose to invest for health, comfort, or environmental reasons even if the financial benefits are low.⁶

We hypothesized that perhaps the presence of a payback period metric was indicative of having numerous metrics and that the presence of multiple metrics was what affected overall quality (as opposed to payback specifically). This was partly supported. The total number of metrics was borderline significantly correlated with overall quality, $r = .30$, $p = .05$.

The second element that significantly predicted experts' ratings of overall quality was the presence of some type of social comparison information in the reports, such as "Your home compared to the average home" or "Your home compared to similar homes." Reports with social comparison information were rated higher ($M = 4.72$) than reports without this information ($M = 3.72$), $t(40) = 2.64$, $p = .01$. In their comments, our experts often made statements such as "Excellent comparison to other homes" for reports that had a comparison. Reports that did not have social comparison information often received comments like "No evidence of integration of principles from behavioral science. There is no mention of norms, what others are doing, or comparisons." Notably, customers who

⁶ Some programs are reducing the barrier of long payback periods by offering financial solutions. Energy as a service (EaaS) is a model that has worked in commercial and industrial markets by offering energy upgrades that are paid for over time as they save energy and money. SEALED is an EaaS provider that works in the residential sector, in some cases helping homeowners invest in upgrades with long payback periods.

received reports rated social comparisons as relatively unimportant in assessment reports (this is discussed in more detail later in this paper).

EXPERTS' COMMENTS ON STRENGTHS AND WEAKNESSES OF REPORTS

Experts independently reviewed each report and provided rich and detailed commentary. We coded and summarized these comments with the aid of qualitative data analysis software. Those results are presented in Appendix B. Experts also participated in focus groups, during which they discussed their most important findings. In general, reviewers had more to say about weaknesses than strengths of the reports. In their independent reviews, they used an average of 30 words to describe strengths and an average of 59 words to describe weaknesses). This was echoed in the focus group discussions.

General Agreement among Experts

Experts tended to describe the strengths and weaknesses of the reports along similar dimensions. Strengths, when they were found, were the inverse of weaknesses that were described in other reports. In general, all three types of experts (assessors, graphic designers, and behavioral scientists) tended to agree on the strengths and weaknesses of the assessment reports. These centered primarily on the summary and recommendations sections.

The focus groups identified seven common themes mentioned by all three groups of experts:

- Most reports could be considerably improved.
- Personalization and custom tailoring of reports is key.
- Reports should have a prioritized list of recommendations and a summary up front, followed by thorough details in the back that allow engaged customers to learn more and find the information they need for following through.
- Reports should avoid being too long and should not overwhelm people with information that they do not need (e.g., an extensive list of all possible replacement options, as opposed to a small number of options that address specific customer concerns).
- Reports should avoid technical language.
- Credibility is gained by tailoring recommendations and avoiding perceived conflicts of interest. For instance, suggesting options that do not earn the assessors any money, and acknowledging potential conflicts, can help build credibility and overcome concerns when, say, an HVAC firm finds that HVAC upgrades are needed.
- Reports should include multiple metrics, not just financial or energy savings, and in particular should include metrics that matter to customers.

Specific Comments by Experts, Based on Areas of Expertise

Reviewers also made some comments based on their specific areas of expertise. Behavioral scientists were more likely to comment on use of behavioral insights (e.g., social norms); location of recommendations within the report; use of language; lack of personalization;

confusing graphics, charts, and tables; and the inclusion of photos and thermal images. Graphic designers and marketers specifically commented on problems with design and formatting, fonts, and the inclusion of unnecessary information. Assessors mentioned concerns with how the assessments were conducted, how results were presented, and the lack of mention of nonfinancial benefits.

Home energy assessors noted:

- Results must be presented in person.
- Some assessments appeared to be poorly conducted.
- Payback periods and savings were sometimes exaggerated, and the reports could have focused on other benefits instead.
- It is problematic that most assessments are provided by utilities that are usually pushing specific recommendations.⁷

Behavioral scientists noted:

- Removing behavioral barriers is key, particularly by providing specific information on how to take next steps.
- Behavioral science should inform both the report writing and in-person assessment in terms of choosing which metrics to use, soliciting commitment during the assessment, actively engaging the homeowner, explaining “why” rather than just “what,” creating a story or narrative, reducing homeowners’ uncertainty, using vivid language, and using heuristics.
- The summary and recommendations should use a direct, personal narrative style.

Graphic designers noted:

- Good (or bad) design can affect credibility.
- Photos and graphs are good, but they need to be high quality and clear and have explanations or notations (e.g., arrows, circles, and highlights). Also, standard photos should be placed next to thermal images.
- Some reports used good design principles (e.g., good color scheme including the use of blue to create calmness and clarity; good branding; and strong visual hierarchy with dividers, labeling, subheaders, and so on).

Eye-Tracking Study

Eye tracking is frequently used by marketing companies to test their websites or advertising materials with customers. We condensed three of the highest rated and most visually interesting reports to five pages each and included them in an eye-tracking study using a nationally representative sample of homeowners ($n = 450$). Each homeowner viewed one of the three reports. Eye tracking services were provided by EyeSee, a company specializing in online eye-tracking research using proprietary software. The EyeSee platform allows

⁷ This was a general comment about assessments, not specific to reports.

researchers to present visual stimuli to online survey participants while tracking their eye movements using their own webcams. In this way, participants working from their personal computers could complete a survey and read actual home energy reports while their eye movements were observed. This allowed us to determine which elements of the reports caught and kept attention and how the design and content of the reports influenced participants' patterns of viewing. By recruiting a nationally representative sample of homeowners, we were also able to use the survey to corroborate some of the experts' ideas on reports that were elicited in the previous part of the project.⁸

Each respondent read a report randomly selected from the three we had abridged and then completed a survey about that report. Complete procedures for participant recruitment, report selection, and survey design are presented in Appendix A.

WHAT HOMEOWNERS LOOKED AT IN THE REPORTS

We examined the eye gaze information of homeowners reading reports using visual maps overlaid on the original assessment report documents. We specifically examined four outcome measures: scan path, frequency of fixation on different elements, gaze duration (time on each element), and the percentage of participants that paid attention to each element. We present a few examples in this section; complete visualizations of outcomes from all three reports are provided in Appendix C.

Scan Path

Scan path is shown in figure 6 as the average gaze starting point and direction across all viewers. All else being equal, we found that homeowners' gazes tended to start in the top third of the page, zigzag around to view different elements, and then gradually move down to the bottom of the page as they finished their scan. Pages with large images or with content that filled only the top half of the page sometimes changed this pattern by discouraging viewers from moving their gaze around the whole document. Although viewers generally scanned the whole page and attempted to view every block of text, diagram, graph, and photo, they stuck to the top-third-to-bottom pattern, based on the position of elements rather than interestingness or importance. Elements at the bottom of the page were always viewed last. We also noticed that when pages had many elements, gaze patterns omitted some of them. It appears that "busy" pages may make it hard for participants to move their gazes to each element and thus led them to skip some.

⁸ Our sample reflected most of the characteristics of American homeowners, except that they had slightly higher levels of education. Appendix A describes our sample's characteristics as compared with the national average.

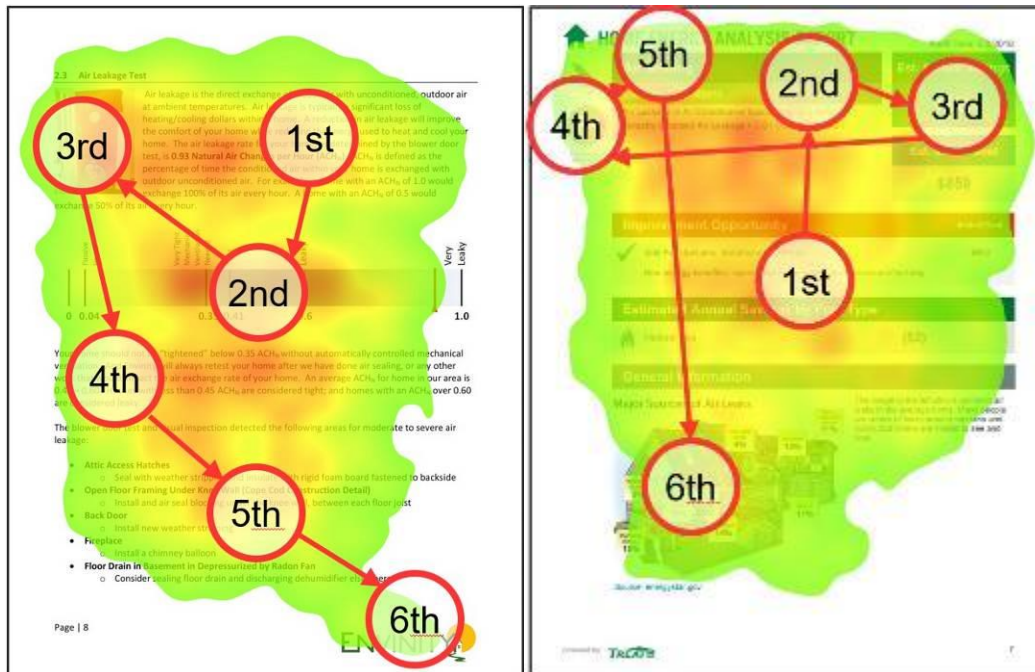


Figure 6. Examples of scan path (numbers and arrows) and frequency of fixations (red shading indicating more fixations). The page on the left has a more natural reading flow and, as such, is less likely to have readers who skip sections.

Frequency of Fixation on Different Elements

Frequency of fixations can be seen in figure 6 as increasing intensity of the red shading. In all three reports, we found that blocks of text were the elements that received the most fixations. This is likely because they contained the most information, and participants were instructed to read and recall the information as though it pertained to their own homes. Relative to the density of text, we were surprised to find that summary tables with information about results and recommendations tended to receive many fixations, as did diagrams and graphs. We included thermal images in only one five-page version of the reports, and although they did not appear to receive significantly more fixations than other images, the text above the thermal images did. This preliminary observation may indicate that thermal images draw attention to text describing the efficiency issue being depicted. However we would need more data to support (or reject) this hypothesis.

Gaze Duration (Time Spent Looking at Each Element)

Gaze duration was measured as the time homeowners spent looking at various predetermined areas of interest. Overall, gaze duration tended to be longer on larger items, items earlier in the report, and items higher up on each page. Text blocks generally held gazes longer than images did, especially narrative summaries of results in which assessors wrote clearly and directly to the homeowner using a first-person voice. Homeowners spent particularly long periods reading summary tables of results and recommendations. For example (figure 7), in the H.E.A.T. Squad report, tables making up the third page had fewer words than text blocks on the second page, yet readers looked at them longer (32.6 seconds looking at tables on page 3 versus 27.6 seconds looking at text on page 2). Thus, homeowners read more slowly when examining summary tables in the H.E.A.T. Squad report than when reading other text blocks (in this case, 14.42 words per second versus 7.55

words per second), despite the fact that later pages usually received less attention.⁹ Gaze duration could indicate confusion, but it is usually a sign of interest in the item being viewed (Goldberg and Wichansky 2003). In the Envinity report, the image with the longest viewing time was the thermal image, and the text above the image received longer viewing times than would be expected for the last page of a report. Again, this is a possible indication that thermal images draw interest, but further research on this topic is necessary.



Figure 7. Example of gaze duration (number of seconds in each area) and percentage of homeowners who viewed each element. Note that tables are often read for longer than paragraphs despite having fewer words (right image), and that when photos are placed to the left of text, those photos are more likely to be seen (left image). Also note that elements toward the bottoms of pages are less likely to be viewed.

Percentage of Homeowners Seeing Each Element

Elements closer to the start of each report and around the top to the middle third of each page are most likely to be seen by participants. Elements at the bottom of each page are skipped by more readers (e.g., 13–21% of H.E.A.T. Squad report readers skipped text or tables at the bottoms of pages two and three). Notably, however, pages that have many competing elements, even if they are on the first page, may have some elements that are skipped. Experts at EyeSee also note that when images are placed to the left of text blocks, they are more likely to be seen than when they are placed to the right (as demonstrated in figure 7). This could be because English is read from left to right. Nevertheless, text blocks with important diagnostic information were viewed by nearly all readers (at least 79% in each case).

⁹ Despite instructions that survey participants would be quizzed and that they should therefore read the reports carefully (as if the results pertained to their own homes), they nevertheless read them at faster than average reading speeds.

Attention and Viewing Takeaways

In terms of text and images, we concluded:

- A mix of images and text helps catch, guide, and maintain attention.
- Image and text objects both get attention, but too many images can result in diffuse focus, and some images not getting enough attention.
- Simple graphs and diagrams catch attention.
- Homeowners are drawn to all elements in reports (text blocks, images, headings, graphs, and tables), and no particular type of element necessarily attracts more attention than others.
- More evidence is needed, but thermal images appear to hold attention for a long time and may encourage readers to read the accompanying text.
- Homeowners focus attention on text and spend most of their time reading as opposed to examining images (possibly partly because they have been instructed to do so). They tend to view all the text areas containing critical diagnostic information and recommendations.
- Narrative summaries and results tables draw and hold attention.

We also came to some conclusions about the structure of the report:

- Attention wanes from page one to page five and from top to bottom of each page (indicated by time spent on each element and percentage of people skipping sections). A good solution might be to put more-important or denser information at the start and at the tops of pages (if it is critical for the homeowner) and to put pictures or less-dense paragraphs toward the end of the report, or provide only half a page of information for later pages.
- Avoiding clutter improves attention. Reports that have only about four to five well-spaced elements (e.g., text blocks, diagrams, photos, graphs) per page tend to discourage viewers from skipping sections.

SURVEY OF EYE-TRACKING STUDY PARTICIPANTS

In addition to observing homeowners' eye movements while they read an assessment report, we asked participants to complete a survey about the report they viewed. After each homeowner had viewed one of the three condensed reports, we asked them content recall questions and a question about the number of recommendations they would follow if they received the report for their own homes. We also asked them to rate the quality of the summary and recommendations sections, the overall understandability, the overall interestingness and appeal, and the complete overall rating of the report. In two open-ended questions, we also asked participants what they liked and disliked about the reports.

There was no statistically significant difference among the participants' survey responses to the three reports on most questions. Content recall, overall quality rating, interestingness and appeal, understandability, and quality of the summary and recommendations sections did not significantly differ between reports (all $ps > .05$). Homeowners scored an average of

6.2 out of 9 on content recall questions. They gave the reports an average score of 79 out of 100 on overall quality, 74 on interestingness and appeal, 77 on understandability, and 79 on the quality of the summary and recommendations sections.

The willingness to invest in recommended upgrades differed by report, $F(2, 447) = 11.24, p < .001$, likely because the recommendations were different and had varying costs and benefits. The abridged report that elicited the highest willingness to invest recommended only three upgrades and highlighted one recommendation that happened to have high annual savings and relatively low cost. It also noted multiple nonenergy benefits and included a good narrative summary. Thus, we cannot be certain why the willingness to upgrade was significantly higher for this report than for the other two, but it could have to do with costs, benefits, message framing, or presentation.

Since these were reports that experts rated highly and that we reduced to five impactful pages, we were not surprised that participants generally liked the reports. The most common aspects they said they liked were these:

- The reports provided specific recommendations and tips for energy efficiency (35%).
- They provided a great breakdown of costs and savings (31%).
- They organized information in sections (30%).
- They were very informative and educational (29%).
- They were concise, comprehensible, easy to read, and easy to understand (25%).

When asked what they did not like about the reports, the largest proportion of homeowners said there was nothing they disliked (47%). The most common answers that others gave were:

- The reports were too technical or hard to understand (10%).
- The recommendations had high costs relative to savings (10%).
- Overall layout and concept of the report was poor (9%).
- The reports were too long to read (8%).¹⁰

At the end of the study, we offered participants an opportunity to engage in an actual energy efficiency behavior: clicking on a link to an external website that they could use to search for and contact a certified home energy assessor in their region. We were encouraged to see that homeowners who indicated a high willingness to invest in hypothetical recommended upgrades were also more likely to search for an actual home energy assessor, $t(448) = 4.95, p < .001$. Those who clicked the link estimated they were willing to invest in an average of 73% of recommended upgrades, whereas those who did not click the link estimated they were willing to invest in 59% of recommended upgrades. This suggests that although our survey involved reports that were not directly relevant to the homeowners

¹⁰ Positive comments about the reports being concise and negative comments about the reports being too long could each be partly explained by the context in which the reports were read. The reports were abbreviated from their original lengths, and participants may have somewhat strained to read the entire contents of the reports while conducting the eye-tracking task.

who completed the survey, their hypothetical answers about likelihood of action may be a reasonable proxy for at least one type of real-world efficiency behavior, searching for a local residential assessor.

Customer Survey

In the last part of the study, we surveyed a group of homeowners who had received residential energy efficiency assessments. We conducted preliminary interviews with 26 assessment recipients and used the results of those interviews to create a large-scale survey for customers served by two organizations. We received 606 surveys from home energy assessment recipients served by ComEd (northern Illinois, $n = 236$) and NYSEERDA (New York State, $n = 370$). All of the customers had received home assessments within one year of completing the survey. Although these assessment recipients may not be representative of assessment recipients nationally, they do represent a large swath of both urban and rural populations in the US Northeast and Midwest. Additional details regarding sampling methodology and the survey itself can be found in Appendix A.

The purpose of the customer survey was to learn what recipients think of their assessment reports and, more broadly, to learn how they respond to receiving an assessment. By surveying recipients, we sought to learn if customers and experts agree on how reports should be written and if people who received assessments respond the way non-recipient experts believe they should.

SATISFACTION WITH THE ASSESSMENT

On average, assessment recipients rated their satisfaction with the assessment as 78 out of 100. Overall satisfaction was significantly correlated with satisfaction with how the assessment results were delivered ($r = .80, p < .001$), perceived quality of the report ($r = .83, p < .001$), and how generic or customized the respondents perceived the report to be ($r = .50, p < .001$).

THE REPORT

We were surprised to learn that 15% of home energy assessment recipients did not recall receiving written results of their assessments. Across all customers, 5% of respondents stated that they did not receive results at all, and another 10% said they received their results only verbally. In written comments (“other, describe”), 10 assessment recipients made comments about either not getting a report when they wanted one, having to ask repeatedly for it, or having to wait a long time to get it. The form in which participants received the results of their assessments is presented in table 7. The average rating of satisfaction with how results were presented was 79 out of 100, and customers who were satisfied with the form of delivery tended to also be satisfied with the assessment overall ($r = .80, p < .001$). Among participants who received a written report (digital or printed), overall satisfaction with the assessment was significantly higher when they also received an oral explanation of results (86 out of 100) than when they did not (76 out of 100), $t(491) = 503, p < .001$.

Table 7. Form in which customers received assessment results

| Form of results | Frequency | % of customers |
|--|-----------|----------------|
| No results received | 32 | 5% |
| Printed only | 124 | 21% |
| Digital only | 86 | 14% |
| Verbal only (in person or by phone) | 63 | 10% |
| Printed + digital | 18 | 3% |
| Printed + verbal | 106 | 18% |
| Digital + verbal | 62 | 10% |
| Printed + digital + verbal | 36 | 6% |
| Did not answer, I don't know, or "other" | 79 | 13% |
| Total | 606 | 100% |

Personal Tailoring

Experts in our focus groups stressed the importance of tailoring assessment reports to customers' homes and addressing specific concerns. Therefore we asked assessment recipients to rate the level of personalization of their reports from 0 ("none of the information was specific to my home") to 100 ("all of the information was specific to my home"). On average, customers gave their reports a score of 75.

Digital or Printed Reports

We examined whether customers were more satisfied with receiving digital or printed reports. Customers gave both types an average rating of 82 out of 100. Ratings were significantly correlated with the level of customization of the report ($r = .61, p < .001$). The difference between printed and digital reports was nonsignificant regardless of whether customers also received a verbal explanation, $t(163) = .39, p = .7$, or not, $t(200) = .39, p = .7$.

ACTING ON THE REPORT

A primary goal of the customer survey was to ascertain how people respond to receiving assessment reports and what may cause a report to be influential (or not). We asked customers who received recommendations for upgrading what the likelihood was that they would act within the next year, from 0 (extremely unlikely) to 100 (extremely likely or already acted on a recommendation). Overall, respondents rated their likelihood of acting as 76, and answers were significantly correlated with overall satisfaction with the assessment ($r = .4, p < .001$), satisfaction with how results were delivered ($r = .8, p < .001$), quality of the report ($r = .47, p < .001$), and level of customization of the report ($r = .24, p < .001$). Each of these factors is important, but they are also closely related to one another. To learn if any of these predicted likelihood of action on their own, we conducted a regression analysis including all four factors, $F(4, 355) = 28.13, p < .001, R^2 = .24$. We found that, with all factors entered in the model simultaneously, satisfaction with how results were delivered to customers was the only significant unique predictor of likelihood of action, $\beta = .36, p < .001$. Thus, assessors should pay particular attention to how they deliver results to customers because this may influence the likelihood of follow-through.

Explaining Results Orally May Indirectly Lead to Action

Expert assessors told us that delivering results to customers in person is vitally important. This survey found that customers who received oral delivery along with written reports were significantly more satisfied with their assessments overall, so we decided to test if they might also be more satisfied with how the results were delivered to them and, in turn, how likely they were to act. After all, satisfaction with the form of delivery of the results is a strong unique predictor of action intention. We hypothesized that customers who received both oral and written reports would be more satisfied with the form of delivery of their results, which would lead to a greater likelihood of action.

A mediation analysis, available in Appendix D, demonstrated that this was indeed the case. Although there was no direct connection between receiving an oral explanation and likelihood of action, we did find an indirect connection. Customers who received an oral and written report were more satisfied with how results were delivered than those who received only a written report. In turn, this elevated satisfaction led to a higher likelihood of acting.

CUSTOMERS' ACTIONS AFTER RECEIVING ASSESSMENTS

We constructed a list of potential responses to receiving an assessment, based on preliminary interviews with ComEd customers ($n = 26$). The most commonly selected option (from the pooled ComEd and NYSERDA populations) was saving the report for their records (41%), followed by installing the recommended upgrades (34%). Twenty-seven percent of customers applied for a rebate. NYSERDA customers were additionally asked if upgrades were installed with or without the help of a contractor. In that survey, 26% of all customers stated that they had the help of a contractor, and 15% of all customers stated that they installed recommended upgrades without the help of a contractor.

In the NYSERDA survey, we also specifically followed up with customers who stated that they were either extremely likely to follow through on recommendations or had already done so (i.e., answered 100/100 on the question about likelihood of action). Most respondents ($n = 113$, 88%) explained the actions they had taken:¹¹

- 51% installed insulation.
- 35% weather-stripped or air-sealed windows, doors, or attics.
- 29% upgraded or installed a new furnace.
- 19% purchased a new water heater.

NYSERDA customers were also asked if there were any recommendations made by the assessors that they planned not to follow. Thirty-two percent of respondents answered this question, although most provided explanations of *why* they were not going to follow through rather than *what* they planned not to do ($n = 51$), and several used this opportunity to complain about their assessments rather than answer the question that was posed ($n = 10$). Of the remaining participants, the largest proportion explained that they did not plan

¹¹ Respondents could have taken multiple actions.

on installing insulation or weather stripping ($n = 35$), or that they did not plan on upgrading their furnace or HVAC system ($n = 12$). Insulation, weather stripping, and furnace improvements are among the most frequently recommended upgrades, which is why they are on both the most likely and least likely upgrades lists.

We also asked NYSERDA customers to select from lists of reasons why they followed through, or chose not to follow through, on assessors' recommendations. These lists were derived from previous research and consultation with NYSERDA staff. The most commonly selected reasons for taking action (among those who indicated they were extremely likely to act or had already done so), were:

- Comfort (80%)
- Saving energy (70%)
- Saving money (54%)
- Fixing problems in the home (51%)
- Availability of a NYSERDA program (48%)

The most commonly selected reasons for not following through on recommendations (among 206 people who selected at least one reason) were:

- The recommendations were too expensive (62%).
- The customer was uncertain about costs and savings (33%).
- Other home improvements were more important (30%).

These reasons give us a rough idea of potential behavioral motivators (and inhibitors), but they are limited as well. Self-reported motivations can be subjective or unreliable. Sometimes they are biased by what respondents believe is socially desirable, by imperfect memories, or by the difficulty of truly knowing or articulating all unconscious and contextual motivators (such as social norms and emotional factors).

CUSTOMER INPUT FOR CREATING AN IDEAL REPORT

In an open-ended question, both ComEd and NYSERDA customers described what they believed were the most useful parts of the reports they received. The most common answers (across both surveys) were that the most useful parts were:

- Recommendations on how to make better lighting choices to save energy
- Specific, easy recommendations on how to save money and energy
- Information on heat, air, and energy loss
- Generally, the whole report

NYSERDA customers were additionally asked an open-ended question regarding the least useful parts of the report. Most respondents explained that, in fact, they found all the information in the report to be useful ($n = 59$). However a small number indicated:

- The information on cost analysis and savings estimates was not useful ($n = 9$).
- The assessor was incompetent or the overall process was flawed ($n = 9$).
- There were inaccuracies in the assessment or report ($n = 8$).

- Presenting expensive, overdetailed solutions was not useful ($n = 8$).
- The report used largely standard, boilerplate language rather than being tailored to their homes ($n = 8$).

Customer Beliefs Regarding Potentially Important Elements in the Report

We asked customers what elements they believe would be important to include in a home energy assessment report. These customers have experience with reports and therefore are in a good position to inform us about what they actually find most useful. We asked them to rate 11 elements (preselected on the basis of interviews with ComEd customers, previous research, and discussion with experts). Each item could be rated from 0 (unimportant) to 100 (extremely important). As shown in figure 8, most items were rated highly. Information about the building envelope was the highest-rated item (88 out of 100) and neighbor comparison was lowest (59 out of 100).

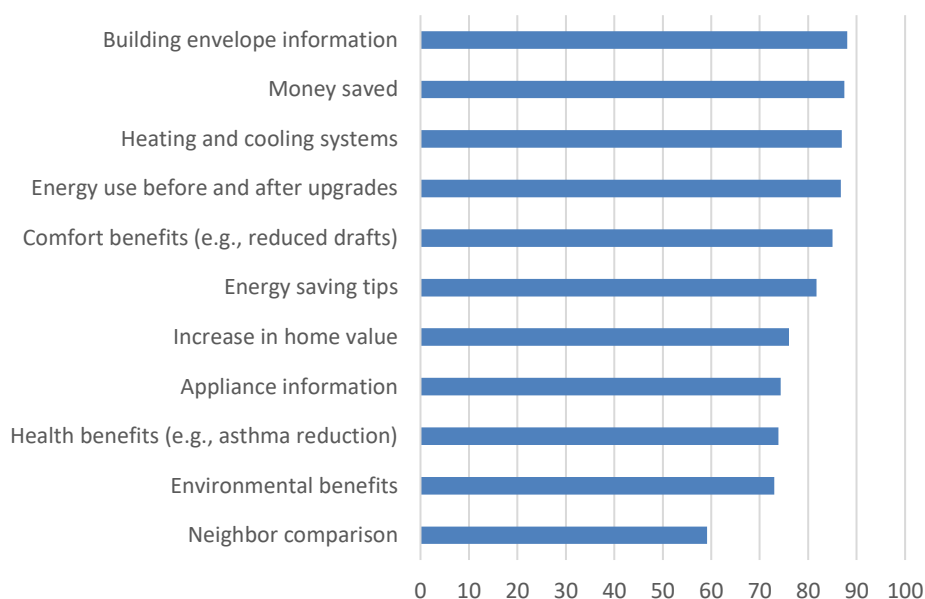


Figure 8. Homeowners' ratings of importance of potential elements of a residential energy efficiency assessment

Neighbor comparison is an interesting element. Although expert reviewers suggested that social comparison information (such as neighbor comparison) was important to have in efficiency assessment reports, we were not surprised that homeowners themselves rated neighbor comparison as relatively unimportant. This is partly because neighbor comparison is only one type of social comparison (comparison to an “average home” is another example), and partly because people typically underestimate and even deny the effects of social norms on their behavior (Cialdini 2007). Thus, although social comparison information is compelling, many people do not like being compared with others, particularly if those others are judged to be better (in this case, more energy efficient). Home energy reports describing social norms of energy use by similar neighbors tend to prompt both habitual energy reduction and some durable home upgrades (Brandon et al. 2017).

TAKEAWAYS FROM CUSTOMER SURVEY

Assessment recipients were generally satisfied with their reports (rating them 82 out of 100) but a little less satisfied with the overall assessment (78 out of 100). When customers followed through with recommended upgrades, they stated doing so for reasons of comfort, energy savings, or cost savings. When they did not follow through, they again gave financial reasons for their decision (too expensive, or not sure about costs and savings).

Based on these findings, we recommend that assessors:

- Ensure that customers actually receive written reports following their assessments
- Deliver the report both in written form and orally; this increases satisfaction with how the results are presented, as well as with the assessment overall. It may also indirectly lead to more action.
- Tailor the report to the home (and homeowner concerns), and avoid boilerplate language when personalized information can be used instead.

Overall Findings and Recommendations

Here are our findings and recommendations for each research question.

WHAT ARE THE MOST COMMON ELEMENTS USED IN RESIDENTIAL ENERGY EFFICIENCY ASSESSMENT REPORTS, AND WHAT ARE THE MOST COMMON RECOMMENDATIONS?

This two-part question can be partly answered through previous research and partly through our coding of sample reports. Sprehn et al. (2015) identified 34 content areas that are common in residential efficiency assessment reports. Of these, the researchers found that four were essential: a summary of the audit procedure, a summary of recommendations, building photographs, and numbers (e.g., kilowatt-hours, dollars, or therms).

In our coding of design, structural, and writing variables, we found that most reports included 1.25 images per page, 56% of which were personally tailored. Images normally occupied less than 25% of the overall report document. Most reports also included graphs (51%) and tables (73%). Only 29% of reports included narrative summaries of findings and recommendations, and only 33% included behavior-change tips. Just over half of the reports (56%) clearly presented next steps for homeowners.

Financial metrics (e.g., annual savings, costs, payback periods) were mentioned in most reports, health inspection results were included in just under half, energy savings metrics were included in about a quarter of the reports, and environmental or water-saving metrics were present in less than one-fifth. The most frequently mentioned benefits of upgrading were cost savings, comfort, and health.

Interviews with home energy assessors conducted in 2010 (Palmer et al. 2013) showed that the most common assessment recommendations were attic insulation (91% recommended fairly often or always), attic or other air sealing (91%), and caulking and sealing of windows or doors (79%). The least common recommendations involved windows (13%); doors (12%); and computers, TVs, or other electronics (10%).

WHAT ARE THE MOST ATTENTION-CATCHING ELEMENTS OF THE REPORTS?

From our eye-tracking study using abridged versions of good reports, we identified a few elements that catch and hold attention. We found that both images and text draw attention, but that too many images can result in some images not getting enough attention. We noticed that participants spent the most time reading text and, in particular, summary tables and narrative summaries of results. Thermal images may help draw and focus attention on issues in the home, but this requires further investigation. Attention normally wanes from the beginning to the end of a report and from the top of each page to the bottom. Therefore important and dense information (if it is critical for the homeowner to know) should be placed at the tops of pages, and pictures or less-dense text and tables should be placed toward the end of the report (with reference materials at the very end). Attention appeared to move more in relation to the structure of a report than in relation to interestingness or importance of the information. Cluttered pages lead to some information being skipped. Therefore including only four to five well-spaced elements (e.g., text blocks, diagrams, photos, graphs) per page, may discourage viewers from skipping sections of the report.

WHICH ELEMENTS OF THE REPORTS ARE MOST PERSUASIVE?

Expert reviews and eye-tracking results provide some indication of the persuasiveness of various elements. The most persuasive elements of a report are those that remove barriers to upgrading, provide clear information, and lend trustworthiness and credibility to the assessment. Generally, persuasiveness results from the report and assessment process as a whole, as opposed to discrete elements of the report alone. For example, presenting the report in both written and verbal form is more persuasive than providing it in only one form or the other. A few elements of the reports themselves were highlighted as being useful. These include:

- An effective text summary near the start of the report, including clearly defined next steps. The summary should be personally tailored and written using a narrative style.
- Summary tables for diagnostics and recommendations
- A mix of images and text blocks, along with simple graphs that clarify information. Images should be of high quality (with attention paid to content, composition, lighting, and so on) and include explanations (e.g., arrows or circles pointing out important information).
- Placing the most important information for the homeowner (the summary of results and recommendations) within the first four to five pages of the report, and placing additional reference information for contractors and interested readers in the back of the report
- Personalization of the report to address homeowner concerns and solutions specifically for the home being assessed (as opposed to general information)
- Showing multiple benefits of upgrading using a variety of metrics (e.g., costs; financial and energy savings; payback periods; and health, comfort, and environmental benefits)
- Simple, clear language – preferably at about an eighth-grade level

- Comparison information, such as how the current home compares with other homes, an average home, or potential future home of the report recipient (whole home or specific parameters such as air leakage, noise, attic, and furnace)
- Prioritized recommendations, with the reason for prioritization clearly explained. This should be based on the customers' concerns and needs and, as such, may change depending on how highly the customer values finances, health, comfort, home value, or other aspects of upgrading.
- Exclusion of information and recommendations that are not relevant to the home being assessed
- Good design that avoids clutter and allows readers' gazes to effortlessly take in each report element

WHAT IS THE PREFERRED LENGTH AND LEVEL OF DETAIL OF A RESIDENTIAL ENERGY EFFICIENCY ASSESSMENT REPORT?

Our expert panel often lamented that reports were too long. However the reports that they rated highly overall did not differ significantly in length from those that they rated poorly. When we delved deeper into comments regarding length, we found that the problem was not the overall length of the report as much as the presence of unnecessary writing and poor report structure. Indeed, some of the longer reports were praised for their thoroughness. The overall length of the report is not as important as excluding information that is irrelevant. For example, reports could exclude rebates that consumers do not qualify for, or they could present fewer options for upgrades (e.g., a small number of excellent options, as opposed to every option available).

The preferred level of detail is difficult to estimate. Expert assessors attest that most homeowners stop reading (or stop paying close attention) to their reports after about the fourth or fifth page. Therefore crucial information should be present in those early pages. However some homeowners will be curious to know more or will desire details about specific upgrades, for their own knowledge or to show to contractors. Having additional information in an appendix or back section could satisfy homeowners at all levels of engagement.

WHICH METRICS DO RECIPIENTS MOST PREFER FOR DESCRIBING COSTS AND BENEFITS/SAVINGS IN THE REPORTS?

Experts agreed that a clear description of up-front costs and savings is essential for the reports, but this project was unable to determine if there was a clear advantage to using certain specific measures (e.g., savings-to-investment ratio or return on investment) or certain timescales (e.g., lifetime savings or annual savings). Reports with payback periods tended to get higher overall ratings from experts than reports without this information because they simplify savings calculations; however, if payback periods are mis-estimated or are quite long (which is common), then they can be very unpersuasive. Assessment recipients themselves rated financial savings and energy use before and after upgrades to be particularly important. More research is needed on how customers perceive, understand, and prefer various metrics.

HOW CAN THE PROCESS OF PRESENTING THE REPORT BE IMPROVED?

As described in Sussman and Chikumbo (2017), numerous interpersonal factors can influence recipients' perceptions of the assessor, perceptions of the assessment, and likelihood of action. For example, assessors perceived as trustworthy, credible, and likable are most apt to be persuasive (Perloff 2003). Listening to homeowner concerns, empathizing, and addressing those concerns in the assessment are important for increasing trust and credibility. Assessors who actively engage homeowners during the assessment, use vivid language, employ message-framing strategies, and elicit some level of commitment to act may also increase their chances of converting an assessment into action (Gonzales, Aronson, and Constanzo 1988).

In our research, we also found additional support for the idea that reports should be presented in person (verbally) as well as in written form. Comments from assessors we spoke to corroborated findings from our survey of assessment recipients. Customers are more likely to be satisfied with the form in which they receive the results, and with their assessments overall, if they receive both verbal and written explanations. This increase in satisfaction appears to lead to increased likelihood of action. Notably, in our sample of assessment recipients, 5% did not recall receiving any results (written or oral) from their assessments, and an additional 10% received their results only in person or over the phone. Undoubtedly this reduces the likelihood of action in some cases.

HOW DO HOMEOWNERS USE THE REPORTS?

On average, customers whom we surveyed who received efficiency assessment reports rated their likelihood of following through with at least one recommendation as 76% (100% indicating extreme likelihood or action already taken). This correlated with overall satisfaction with the assessment, satisfaction with how results were presented, quality of the report, and level of personalization.

Most often, customers saved the reports for their records (41%) and/or installed upgrades (34%), usually with the help of a contractor. About a quarter of participants (27%) said they applied for a rebate, and 14% contacted a contractor for more information.

In terms of specific upgrades, customers who took action most often stated that they installed insulation (51%); weather-stripped or air-sealed windows, doors, or attics (35%); upgraded or installed a furnace (29%); or purchased a new water heater (19%). Conversely, some of these same upgrades were also commonly mentioned as examples of work that report recipients would *not* do (adding insulation or weather stripping or upgrading a furnace or HVAC system). Notably, these results are based on customer surveys from the NYSERDA and ComEd territories of New York State and northern Illinois, respectively. They may differ from results produced in other regions.

Discussion

Residential efficiency assessments are an effective means by which to encourage homeowners to upgrade and renovate existing homes and, as such, are important to any overall energy reduction or climate action strategy. Creating informative and persuasive reports and presenting them effectively to homeowners can further improve these assessment programs.

This study is the most comprehensive examination of residential efficiency assessment reports to date. We used a mixed-method approach to examine sample reports, engaging reviews by experts, homeowners, and recipients of actual assessments. These approaches allowed us to triangulate our findings and posit recommendations that we are confident could improve reports. Future studies could use these results and build on this research.

LIMITATIONS AND FUTURE DIRECTIONS

Although we gathered more assessment reports than other research teams had previously gathered and evaluated them with input from larger and more diverse groups (experts, nationally representative homeowners, and assessment customers), we nevertheless acknowledge that evaluations can be subjective. Given this subjectivity, we suggest further field research testing the application of our findings. The next logical step is to apply the design and content recommendations from this paper to create improved report writing guides and lessons. Field-testing new reports based on these recommendations would allow us to learn how these new reports perform relative to current practices.

CONCLUSIONS

To maximize the persuasiveness of efficiency reports and encourage homeowners to follow through on recommendations, the reports should be well designed, written clearly and succinctly, presented in person, and informed by social science insights. Simplifying the language and reducing unnecessary content can be tricky but go a long way toward improving understandability and the likelihood of future action. The goal of the assessment report is to inform the homeowner, increase credibility and trustworthiness of the assessment, reduce barriers to taking action, and motivate the customer. To do this, we recommend a few improvements:

- Provide key information in the first four to five pages, followed by more detailed, supplementary material at the end.
- Summarize results and recommendations in tables and personalized narrative summaries.
- Prioritize recommendations according to customer concerns and needs, with explanations of the prioritization.
- Maintain a simple and consistent structure with high-quality images that are explained with notations and accompanying text.

Effective reports reduce uncertainty and barriers to action. From a behavioral perspective, assessors could take advantage of key social science insights to improve report persuasiveness. These include:

- Highlighting social norms or providing comparison information, such as comparing the customer's home with the average home and the customer's potential future home.
- Using vivid language, such as metaphors, that brings the report to life.
- Presenting a variety of benefits of upgrading, such as health, comfort, and cost savings.

References

- Aksoezen, M., M. Daniel, U. Hassler, and N. Kohler. 2015. "Building Age as an Indicator for Energy Consumption." *Energy and Buildings* 87: 74–86.
www.sciencedirect.com/science/article/pii/S0378778814009207.
- Alberini, A., and C. Towe. 2015. "Information v. Energy Efficiency Incentives: Evidence from Residential Electricity Consumption in Maryland." *Energy Economics* 52: S30–40.
www.sciencedirect.com/science/article/pii/S0140988315002364.
- Baron, R., and D. Kenny. 1986. "The Moderator-Mediator Variable Distinction in Social Psychological Research: Conceptual, Strategic, and Statistical Considerations." *Journal of Personality and Social Psychology* 51 (6): 1173–82.
www.ncbi.nlm.nih.gov/pubmed/3806354.
- Billingsley, M., C. Stratton, and E. Fadrhonc. 2016. *Energy Advisors: Improving Customer Experience and Efficiency Program Outcomes*. Prepared by Berkeley Lab. Washington, DC: DOE. eta.lbl.gov/publications/energy-advisors-improving-customer.
- Bindemann, M., M. Burton, I. Hooge, R. Jenkins, and E. De Haan. 2005. "Faces Retain Attention." *Psychonomic Bulletin & Review* 12 (6): 1048-53.
link.springer.com/content/pdf/10.3758/BF03206442.pdf.
- Brandon, A., P. Ferraro, J. List, R. Metcalfe, M. Price, and F. Rundhammer. 2017. *Do the Effects of Social Nudges Persist? Theory and Evidence from 38 Natural Field Experiments*. Working Paper No. 23277. Cambridge, MA: NBER (National Bureau of Economic Research). www.nber.org/papers/w23277.
- Breukers, S., E. Heiskanen, R. Mourik, D. Bauknecht, M. Hodson, Y. Barabanova, B. Brohmann, V. Bürger, C. Feenstra, M. Jalas, M. Johnson, P. Maier, S. Marvin, H. Meinel, J. Pariag, M. Rask, S. Rinne, S. Robinson, M. Saastamoinen, J. Salminen, I. Valuntiené, and E. Vadovics. 2009. *Interaction Schemes for Successful Energy Demand Side Management. Building Blocks for a Practicable and Conceptual Framework: Deliverable 5 of the Changing Behaviour Project*. Brussels: European Commission.
www.researchgate.net/publication/330848191_Interaction_schemes_for_successful_ene_rgy_demand_side_management_building_blocks_for_a_practicable_and_conceptual_framework_Deliverable_5.
- Census Bureau. 2017. "American Housing Survey Data: 2015." www.census.gov/programs-surveys/ahs/data.2015.html.
- Cialdini, R. 2007. "Descriptive Social Norms as Underappreciated Sources of Social Control." *Psychometrika* 72 (2): 263. link.springer.com/article/10.1007/s11336-006-1560-6.
- Delmas, M., M. Fischlein, and O. Asensio. 2013. "Information Strategies and Energy Conservation Behavior: A Meta-Analysis of Experimental Studies from 1975 to 2012." *Energy Policy* 61: 729–39.
www.sciencedirect.com/science/article/pii/S0301421513004643.

- DOE (Department of Energy). 2019a. "Partner Resources." Accessed January. betterbuildingsinitiative.energy.gov/home-energy-score/homeenergy-score-partners-partner-resources.
- . 2019b. "Weatherization Assistance Program." Retrieved January. www.energy.gov/eere/wipo/weatherization-assistance-program.
- EIA (Energy Information Administration). 2018a. "Electricity Explained: Use of Electricity." www.eia.gov/energyexplained/index.php?page=electricity_use.
- . 2018b. "Table HC2.3 Structural and Geographic Characteristics of U.S. Homes by Year of Construction, 2015." www.eia.gov/consumption/residential/data/2015/hc/php/hc2.3.php
- Frondel, M., and C. Vance. 2013. "Heterogeneity in the Effect of Home Energy Audits: Theory and Evidence." *Environmental and Resource Economics* 55 (3): 407–18. link.springer.com/article/10.1007/s10640-013-9632-4.
- Fuller, M., C. Kunkel, M. Zimring, I. Hoffman, K. Soroye, and C. Goldman. 2010. *Driving Demand for Home Improvements: Motivating Residential Customers to Invest in Comprehensive Upgrades That Eliminate Energy Waste, Avoid High Bills, and Spur the Economy*. Prepared by Berkeley Lab. Washington, DC: DOE. emp.lbl.gov/sites/all/files/report-low-res-bnl-3960e.pdf.
- Gardner, G., and P. Stern. 2008. "The Short List: The Most Effective Actions US Households Can Take to Curb Climate Change." *Environment: Science and Policy for Sustainable Development* 50 (5): 12–25. www.tandfonline.com/doi/abs/10.3200/ENVT.50.5.12-25.
- Gillingham, K., and T. Tsvetanov. 2018. "Nudging Energy Efficiency Audits: Evidence from a Field Experiment." *Journal of Environmental Economics and Management* 90: 303–16. www.sciencedirect.com/science/article/abs/pii/S0095069617305831.
- Goldberg, J., and A. Wichansky. 2003. "Eye Tracking in Usability Evaluation: A Practitioner's Guide." *The Mind's Eye: Cognitive and Applied Aspects of Eye Movement Research*, J. Hyönä, R. Radach, and H. Deubel, eds., 493–516. Amsterdam: North Holland Publishing Company.
- Gonzales, M., E. Aronson, and M. Costanzo. 1988. "Using Social Cognition and Persuasion to Promote Energy Conservation: A Quasi-Experiment." *Journal of Applied Social Psychology* 18 (12): 1049–66. onlinelibrary.wiley.com/doi/abs/10.1111/j.1559-1816.1988.tb01192.x.
- Hayes, A. 2009. "Beyond Baron and Kenny: Statistical Mediation Analysis in the New Millennium." *Communication Monographs* 76: 408–20. nca.tandfonline.com/doi/abs/10.1080/03637750903310360?journalCode=rcmm20#.XJzEJNKjOQ.

- . 2013. *Introduction to Mediation, Moderation, and Conditional Process Analysis: A Regression-Based Approach*. New York: Guilford Press.
- Holladay, J., J. LaRiviere, D. Novgorodsky, and M. Price. 2016. *Asymmetric Effects of Non-Pecuniary Signals on Search and Purchase Behavior for Energy-Efficient Durable Goods*. NBER Working Paper No. 22939. Cambridge, MA: NBER. www.nber.org/papers/w22939.
- Ingle, A., M. Moezzi, L. Lutzenhiser, and R. Diamond. 2012. "How Well Do Home Energy Audits Serve the Homeowner?" In *Proceedings of the 2012 ACEEE Summer Study on Energy Efficiency in Buildings 2*: 217–229. Washington, DC: ACEEE. www.pdx.edu/cus/sites/www.pdx.edu.cus/files/Ingle%20et%20al%20%282012%29%20How%20Well%20Do%20Home%20Energy%20Audits%20Serve%20the%20Homeowner%3F.pdf.
- Jaffe, A., and R. Stavins. 1994. "The Energy-Efficiency Gap: What Does It Mean?" *Energy Policy* 22 (10): 804–10. www.sciencedirect.com/science/article/pii/0301421594901384.
- Jewell, M. 2014. *Selling Energy: Inspiring Ideas That Get More Projects Approved!* San Francisco: Energy Efficiency Funding Group.
- Johns, J., and T. Wheat. 1984. "Newspaper Readability: Two Crucial Factors." *Journal of Reading* 27 (5): 432–4. www.jstor.org/stable/40032569.
- Ketchman, K., V. Khanna, D. Riley, and M. Bilec. 2016. "Evaluation of a Holistic Energy Assessment Program." *Procedia Engineering* 145: 468–75. www.sciencedirect.com/science/article/pii/S1877705816300248.
- Kirby, S., A. Guin, L. Langham, and A. Chilcote. 2014. "Exploring the Impact of the E-Conservation Residential Energy Audit Program." *Housing and Society* 41 (1): 71–88. www.tandfonline.com/doi/abs/10.1080/08882746.2014.11430622.
- Lawshe, C. 1975. "A Quantitative Approach to Content Validity 1." *Personnel Psychology* 28 (4): 563–75. onlinelibrary.wiley.com/doi/abs/10.1111/j.1744-6570.1975.tb01393.x.
- MacKinnon, D., C. Lockwood, and J. Williams. 2004. "Confidence Limits for the Indirect Effect: Distribution of the Product and Resampling Methods." *Multivariate Behavioral Research* 39 (1): 99–128.
- Murphy, L. 2014. "The Influence of Energy Audits on the Energy Efficiency Investments of Private Owner-Occupied Households in the Netherlands." *Energy Policy* 65: 398–407. www.sciencedirect.com/science/article/pii/S0301421513010380.
- NCES (National Center for Educational Statistics). 2003. "National Assessment of Adult Literacy (NAAL)." nces.ed.gov/naal/kf_demographics.asp.
- NREL (National Renewable Energy Laboratory). 2017. *ResStock – Targeting Energy and Cost Savings for U.S. Homes*. www.nrel.gov/docs/fy17osti/68653.pdf.

- Ojha, P., A. Ismail, and K. Kuppusamy. 2018. "Perusal of Readability with Focus on Web Content Understandability." *Journal of King Saud University – Computer and Information Sciences*. www.sciencedirect.com/science/article/pii/S1319157817303622.
- Palmer, K., M. Walls, H. Gordon, and T. Gerarden. 2013. "Assessing the Energy-Efficiency Information Gap: Results from a Survey of Home Energy Auditors." *Energy Efficiency* 6 (2): 271–92. link.springer.com/article/10.1007/s12053-012-9178-2.
- Palmer, K., M. Walls, and L. O’Keeffe. 2015. *Putting Information into Action: What Explains Follow-Up on Home Energy Audits?* Washington, DC: Resources for the Future. www.rff.org/publications/working-papers/putting-information-into-action-what-explains-follow-up-on-home-energy-audits/.
- Perloff, R. 2003. *The Dynamics of Persuasion: Communication and Attitudes in the 21st Century*. Mahwah, NJ: Lawrence Erlbaum Associates. psycnet.apa.org/record/2003-04019-000.
- Readable. 2019. "How's Your Readability?" www.readable.com.
- Riley, D., M. Whelton, N. Mohr, and S. Barsom. 2012. "An Analysis of Home Energy Assessments." State College: Pennsylvania State University. Shrouf, P., and N. Bolger. 2002. "Mediation in Experimental and Nonexperimental Studies: New Procedures and Recommendations." *Psychological Methods* 7 (4): 422–45. www.ncbi.nlm.nih.gov/pubmed/12530702.
- Sprehn, K., M. Whelton, D. Riley, and G. Okudan-Kremer. 2015. "Content Analysis of Home Energy Audit Reports: Quantitative Methodology for Standardization." *Journal of Architectural Engineering* 21 (4). [ascelibrary.org/doi/abs/10.1061/\(ASCE\)AE.1943-5568.0000170](http://ascelibrary.org/doi/abs/10.1061/(ASCE)AE.1943-5568.0000170).
- Stern, P., E. Aronson, J. Darley, D. Hill, E. Hirst, W. Kempton, and T. Wilbanks. 1986. "The Effectiveness of Incentives for Residential Energy Conservation." *Evaluation Review* 10 (2): 147–76. journals.sagepub.com/doi/abs/10.1177/0193841X8601000201.
- Sussman, R., and M. Chikumbo. 2017. *How to Talk about Home Energy Upgrades*. Washington, DC: ACEEE. aceee.org/research-report/b1701.
- Wefelmeyer, E., and M. Backus. 2017. "Strategies for Using Data Analytics in Testing the Readability Levels of Textbooks: It’s Time to Get Serious." *Procedia Computer Science* 118: 95–9. www.sciencedirect.com/science/article/pii/S1877050917323529.
- Wong-Parodi, G., W. Bruine de Bruin, and C. Canfield. 2013. "Effects of Simplifying Outreach Materials for Energy Conservation Programs That Target Low-Income Consumers." *Energy Policy* 62: 1157–64. www.sciencedirect.com/science/article/pii/S030142151300712X.

Appendix A. Methodology

This project had two parts. The first involved collecting a sample of home energy efficiency assessment reports, coding them, calculating readability metrics for them, submitting them to expert review by individuals and in focus groups, and submitting a subsample to homeowners for study with an eye tracker. The second part involved surveying homeowners who had actually received assessments to understand their opinions and responses to their reports.

CONTENT ANALYSIS

Collecting Reports

We collected sample assessment reports using three strategies. We initially gathered 68 reports, and after removing duplicates and nontraditional reports, we were left with 45. Twenty-four were sent to us in response to a request on the ACEEE website and social media platforms, 10 were publicly available on the Internet (found using a Google search), and 11 were sent to us from sponsors of this project.

Coding Reports

The research team developed a list of potential elements in the report to code, based on previous research, consultation with experts (graphic designers, assessors, and behavioral scientists), and factors that were hypothesized to be important. These fell into the following categories: descriptive factors, basic classification, length, images, colors, use of tables, level of language, metrics used, social norms, described benefits, health and safety information, and report recommendations. We also recorded our own preliminary (nonexpert) ratings of each report. In total we coded 65 variables, including readability metrics and ratings.

The coding for all reports was conducted by a primary coder, with a secondary coder validating the procedure by double-coding 20% of the reports. Inter-rater reliability was deemed acceptable across all nominal data ($\alpha = .80$), ordinal data ($\alpha = .6$), and interval data ($\alpha = .97$). Therefore the primary coder's designations were accepted for the entire dataset.

Readability Metrics

Readability is a measure of the education level needed to understand a piece of text. It is usually presented in terms of minimum grade level. We converted each report into a text document and submitted it to a web-based software tool for readability analysis (app.readable.io).¹² This tool provides several empirically validated metrics for readability, including Flesch-Kincaid, Gunning Fog, Colman-Liau, SMOG index, Automated Readability Index, Flesch Reading Ease, CEFR level, IELTS level, Spache Score, and New Dale-Chall Score (see Ojha, Ismail, and Kuppusamy 2018 for a review and explanation of some metrics). These rely to different extents on word count, sentence count, syllable count, number of letters per word, and commonness of words (comparing words in the reports to lists of common words). We recorded the average estimated grade level across all grade

¹² Given that readability scores can fluctuate based on how readability tools interpret the input (Wefelmeyer and Backus, 2017), we found this process of converting each report to a text file to be critical to obtaining an accurate measurement of readability.

estimation metrics for each report (lower grade levels mean the report is easier to understand). We also recorded an overall readability letter grade (A, B, C, or D) that summarizes the readability scores across all metrics. The program also provided the research team with a word count for each report that we later used to calculate metrics such as number of words per page. These readability metrics were entered into the coding database alongside other elements of the report.

EXPERT REVIEW

We recruited eight behavioral scientists, eight home energy assessors, and six graphic designers/marketing experts to review the reports. Experts were chosen on the basis of years of experience and knowledge of home energy behavior. They were recruited through in-person solicitation at conferences (e.g., Home Performance Coalition, Behavioral Science and Policy Association, and Behavior Energy and Climate Change), and referrals. Several assessors who agreed to review reports were winners of the Home Performance with ENERGY STAR Contractor of the Year Award. The behavioral scientists were primarily faculty members in psychology or behavioral economics with numerous peer-reviewed publications regarding aspects of behavior change and message framing. The graphic designers and marketing experts worked with prominent marketing firms on energy- and environment-related campaigns to encourage behavior change. We paid reviewers a small honorarium for their participation.

Each report was read by three experts, including one of each type (assessor, behavioral scientist, and graphic designer/marketer).¹³ Each reviewer carefully read five to seven reports. Three of the 45 reports, those with the lowest preliminary (nonexpert) overall scores, were not sent for expert review because we did not want to overburden the review panel. The experts completed a scoring sheet asking for numerical scores (from 1 to 7) for quality of summary and recommendations, understandability, persuasiveness, interestingness and appeal, and overall rating. The scoring forms also asked open-ended questions regarding the strengths, weaknesses, and overall impressions of the reports. Finally, reviewers were asked to comment on each report specifically from the perspective of their area of expertise.

We conducted quantitative analyses on scoring questions using SPSS, and qualitative analyses on open-ended questions using NVIVO. We looked for patterns in reviews and related them to elements we coded in each report. In particular, we looked for elements or ratings that best predicted overall expert ratings of quality.

Conducting Focus Groups with Experts

After all experts submitted their personal reviews, we convened these experts in three focus groups to discuss the reports. We had one focus group for assessors, one for behavioral scientists, and one for marketers and graphic designers. During the focus groups, we asked experts to summarize their thoughts on the reports by prompting them with several open-

¹³ For technical reasons and to avoid conflicts of interest, three reports were reviewed by only two people each. Two of those were not reviewed by a home energy assessor, and one was not reviewed by a behavioral scientist.

ended questions. Focus group sessions lasted one to two hours. We summarized the results in point form notes.

EYE-TRACKING STUDY

Three reports were selected and shortened for testing with an eye tracker. Reports were ranked on the basis of expert ratings, and those with high overall ratings along with good scores on “interestingness and appeal” were considered for the eye-tracking component. We wanted to include visually interesting documents that were also representative of generally good reports. Ultimately, we chose to include one report each by Envinity, H.E.A.T. Squad, and Wise Home Energy.

Across all the reports ($n = 42$) that we sent to our expert panel, the average rating of overall quality was 3.94/7, and interestingness/appeal was rated at 3.86/7. The reports selected for the eye-tracking study had much higher scores. The Envinity report received the highest overall rating (6.67/7) and very high scores in all of the other categories (6.17/7 average across all other categories), including 6.33/7 on interestingness and appeal. The H.E.A.T. Squad report received the second-highest overall score (6.17/7) and very high ratings in the other categories (6.33/7 average), including 6/7 on interestingness and appeal. The Wise Home Energy report received a high overall score (4.67/7, tied for ninth), high scores across other categories (4.58/7), and a high interestingness and appeal score (5/7). It was also built with a standard software solution that we wanted to be represented in this part of the study. Experts also commented that this report’s visual aspects were one of its strengths. In our internal evaluation of all 42 reports, ACEEE’s (nonexpert) ranking of the report was third-highest in the group.

Limitations of the eye-tracking survey required that length of each report be no more than five pages. Therefore we cut large portions of each (the average report length was 14 pages long). We also chose reports that had several pages with interesting and unique visual aspects as well as good summary sections. Further, we chose reports that all provided similar recommendations to homeowners, facilitating a fairer comparison among them.

The condensed reports included a cover page, a summary page, and two to three pages of recommendations and explanations. In their original form, all three reports included recommendations for at least two specific upgrades, air sealing and insulation, and these were kept for the final five-page versions. Given that we were using a sample of homeowners who would be receiving reports that did not pertain to their own homes, we felt that their satisfaction with the reports and their intentions to upgrade would probably not truly represent the opinions of people who actually received reports about their own homes (a group we surveyed in the last part of this project in order to address this issue). Instead, the eye-tracking survey was used primarily to test aspects of the reports such as what captured attention, what was liked and disliked, and what was recalled after a brief read-through. Notably, cutting the reports to five pages limited their levels of detail and thoroughness. This restricted our ability to generalize the survey participants’ ratings of report quality.

The five-page versions of the three reports each included a mix of words and full-color images. The Wise report included only generic diagrams and drawings, as well as several

custom bar graphs. The H.E.A.T. Squad report had a few diagrams and photos, including several photos with people in them, and two thermal images.¹⁴ The Envinity report included several photos of the home being assessed, several graphs (pie and bar), a diagram, and a few images with people in them.

EyeSee worked with a panel survey company to recruit 450 homeowners for the eye-tracking survey. We aimed to secure a nationally representative sample of homeowners living in single-family dwellings, either detached or connected to one other unit. Ultimately, the sample that EyeSee recruited matched closely with the key demographics of our target population as reported in the 2015 Census (Census Bureau 2017). Except for education level, each demographic category for our sample (age, income, type of home, and geographic region) differed by no more than 10 percentage points from the target population's demographics. However our sample differs more in education level, with a smaller proportion of homeowners being only high school graduates (census = 42.6%, sample = 16.22%, difference = 26.38%), and more of the sample having an associate's degree or some college (census = 9.95%, sample = 32.00%, difference = 22.05%) or a bachelor's degree (census = 23.79%, sample = 36.44%, difference = 12.65%). The demographics of our sample, compared to the 2015 census information, are presented in table A1.

Table A1. Homeowner demographics (for single-family detached houses or houses attached to one unit) in our sample and in the 2015 US census

| Demographic | 2015 census | Sample |
|-----------------------|-------------|--------|
| Gender | | |
| Male | [no info] | 42.4% |
| Female | [no info] | 57.6% |
| Age | | |
| Under 25 years old | 0.63% | 1.33% |
| 25 to 29 years old | 2.82% | 5.11% |
| 30 to 34 years old | 5.71% | 8.22% |
| 35 to 44 years old | 16.03% | 19.78% |
| 45 to 54 years old | 21.77% | 18.89% |
| 55 to 64 years old | 23.91% | 23.56% |
| 65 to 74 years old | 17.20% | 18.44% |
| 75 years old and over | 11.94% | 4.67% |
| Income | | |
| Less than \$20,000 | 11.22% | 3.36% |
| \$20,000 to \$29,999 | 7.65% | 6.95% |

¹⁴ Although we did not find human imagery to have a particularly strong impact in this study, other research suggests that images of people tend to draw more attention than other objects (Bindemann et al. 2005).

| Demographic | 2015 census | Sample |
|--|-------------|---------|
| \$30,000 to \$39,999 | 7.94% | 12.78% |
| \$40,000 to \$49,999 | 7.88% | 8.97% |
| \$50,000 to \$59,999 | 7.65% | 12.11% |
| \$60,000 to \$79,999 | 14.05% | 17.94% |
| \$80,000 to \$99,999 | 11.24% | 14.80% |
| \$100,000 to \$119,999 | 9.10% | 8.97% |
| \$120,000 or more | 23.28% | 14.13% |
| Geographic region | | |
| Northeast | 17.46% | 19.33% |
| Midwest | 24.94% | 27.11% |
| South | 37.15% | 41.56% |
| West | 20.46% | 12.00% |
| Education | | |
| Less than high school | 8.07% | 0.44% |
| High school graduate (or equivalency) | 42.60% | 16.22% |
| Associate's degree | 9.95% | 32.00%* |
| Bachelor's degree | 23.79% | 36.44% |
| Graduate or professional degree | 15.58% | 14.89% |
| Type of home | | |
| Single detached house | 94% | 93% |
| House attached to one other unit | 6% | 7% |

*In our survey, we also included homeowners with some college in this bracket, thus inflating it slightly.

Each respondent was randomly assigned to read one of the three assessment reports and then answered questions about that report. One-third of the sample read the Envinity report, one-third read the H.E.A.T. Squad report, and one-third read the Wise report. In each group, 60% of the respondents read the reports while their eye movements were tracked ($n = 90$), and 40% ($n = 60$) read the reports without being observed. Each report was presented one page at a time, with participants clicking forward to the next page at their own pace. They were instructed to read the report as though it pertained to their own homes, and they were told that their recall of the details of the report would be tested. All participants then completed the survey. There was no significant difference in content recall scores between participants who read the reports while being observed ($M = 6.33/9$) and those that were not ($M = 6.11/9$), $t(448) = 1.77$, $p = .08$. Therefore we assumed there was no systematic difference in how participants read and understood the reports between the two groups. Consequently, we pooled all participants' scores together for all survey analyses.

Based on previous reviews of eye-tracker research (e.g., Goldberg and Wichansky 2003), we chose to examine several common eye-tracking measures (these are mapped on the actual reports in Appendix B):

- Scan path (the direction of gaze movement and order in which elements of the reports were viewed)
- Number of fixations in different areas (the frequency of participants viewing certain non-predetermined areas of the reports, indicated by dark red coloration in heat maps)
- Gaze duration for specific areas of interest (the amount of time that subjects look at predetermined areas of the reports, such as blocks of text or images)
- Number of participants who looked at specific areas of interest (the percentage of report readers who viewed or skipped elements, such as blocks of text or images)

Personalized Feedback for Organizations that Provided Reports

As part of our commitment to making a direct impact on residential efficiency assessments, we created customized evaluations for each organization that provided us with sample reports for this study. These evaluations described specific (anonymous) expert feedback and provided the scoring for the reports that each energy assessment company gave to us. This feedback could help the assessment organizations redesign and improve their reports to better their assessment outcomes.

CUSTOMER SURVEY

To complement our evaluation of sample assessment reports, we also conducted a survey of customers who actually had an assessment and received a report. We gained permission to survey customers who received assessments through ComEd (a utility serving northern Illinois) and the New York State Energy Research and Development Authority (NYSERDA). Each group provided a database of residential customers who had received home energy assessments by one of its contracted providers no more than one year prior to the survey. ComEd customers received assessments from one of two companies, each with a roster of licensed contractors in its service area. NYSERDA customers received assessments conducted by one of 29 campaign partners.

The NYSERDA database contained 12,050 customers, of whom 4,395 had provided email addresses. The ComEd database contained 16,421 customers, all of whom had provided phone numbers and 3,924 of whom had given their email addresses. We began with preliminary phone interviews to 26 ComEd customers to ask open-ended questions regarding their assessment experiences and opinions on their reports.¹⁵ On the basis of these interviews and background literature reviews, we designed and distributed an online survey to the remaining ComEd assessment recipients for whom we had email addresses. This survey was then augmented with several additional questions and sent to all NYSERDA assessment recipients who'd provided their email addresses. To encourage

¹⁵ We called 600 customers in order to complete 26 interviews. Customers who were called were not sent surveys.

participation, we offered to enter interviewees and survey respondents into a drawing for a gift card. Surveys took approximately 5 to 10 minutes to complete.

Primary analyses were conducted on the combined outcomes of the NYSERDA and ComEd customer surveys. These were supplemented by a few analyses of NYSERDA-only data. Quantitative questions were analyzed using SPSS and qualitative questions were analyzed using NVIVO. We examined factors within the survey to learn if any predicted overall satisfaction or self-reported likelihood of action.

Appendix B. Qualitative Analysis of Individual Expert Reviews

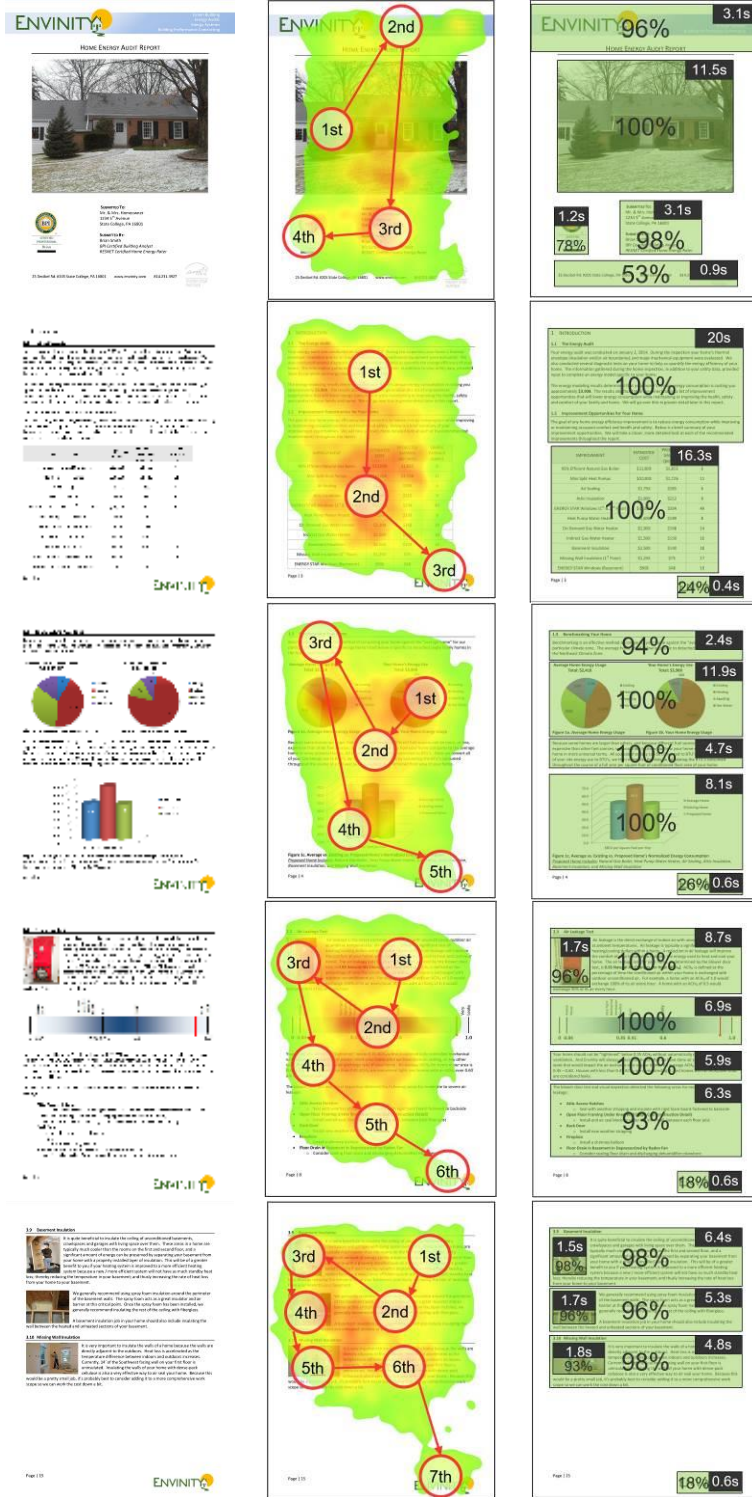
Expert reviewers were asked to first conduct independent evaluations of each assessment report (up to seven were assigned to each expert) and then to come together to discuss the reports in focus groups. The qualitative analysis of experts' independent reviews are presented in table B1.

Table B1. Report strengths and weaknesses

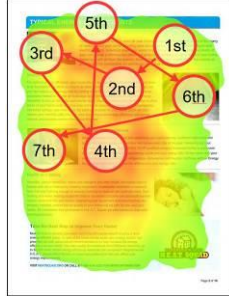
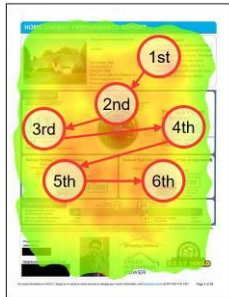
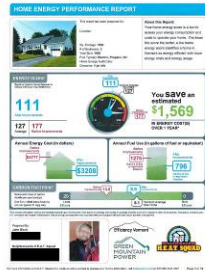
| Characteristic | Strengths | Weaknesses |
|--|---|---|
| Recommendations: <ul style="list-style-type: none"> ● Explanation ● Prioritization ● Ease of finding in report ● Presentation ● Content | <ul style="list-style-type: none"> ● Explanation of recommendations ● Prioritized recommendations ● Upgrades presented in packages or grouped together ● Specific content areas and assessment results ● Up-front summary of findings and takeaways ● Behavior and low- and no-cost tips for saving | <ul style="list-style-type: none"> ● Recommendations that are not well explained ● Recommendations that are not prioritized, or unclear reasons for prioritizing ● Problems with content of recommendations or how assessment was conducted ● Recommendations or important information that are not easy to find ● Recommendations section that is generally poor or not well designed |
| Organization, layout, and formatting | <ul style="list-style-type: none"> ● Layout, organization, and consistent formatting | <ul style="list-style-type: none"> ● Poor organization ● Fonts that are small, inconsistent, poorly colored, or difficult to read |
| Photos | <ul style="list-style-type: none"> ● Good photos and thermal images | <ul style="list-style-type: none"> ● Poor photos or images in general |
| Data visualization and graphics | <ul style="list-style-type: none"> ● Graphics, icons, graphs, charts, tables, and diagrams | <ul style="list-style-type: none"> ● Problems with graphics, charts, tables, graphs, or diagrams ● Too many images, tables, charts, or graphs |
| Personalization | <ul style="list-style-type: none"> ● Personalization and tailoring | <ul style="list-style-type: none"> ● Lack of adequate personalization |
| Design | <ul style="list-style-type: none"> ● Overall design | <ul style="list-style-type: none"> ● Problems with design or formatting |
| Clarity | <ul style="list-style-type: none"> ● Writing that is clear and simple | <ul style="list-style-type: none"> ● Not enough images, or too many words and dense text with lots of details ● Language that is difficult, confusing, too technical, written for experts, or just bad ● Lack of good summary, introduction, or narrative |
| Information quality and quantity | <ul style="list-style-type: none"> ● Thoroughness ● Rebates, financing, and incentives ● Presentation of costs and/or savings ● General information and overall content ● Diagnosis of energy issues | <ul style="list-style-type: none"> ● Lack of key information, such as cost or rebate info, or content that is not informative enough |
| Facilitating next steps | <ul style="list-style-type: none"> ● Information for taking next steps | <ul style="list-style-type: none"> ● Lack of information required for next steps |
| Context for comparison | <ul style="list-style-type: none"> ● Comparison with other homes or future home | <ul style="list-style-type: none"> ● No context such as comparison to other homes, typical homes, or future home |
| Nonenergy benefits | <ul style="list-style-type: none"> ● Inclusion of nonenergy, nonfinancial benefits | <ul style="list-style-type: none"> ● No mention of benefits beyond energy or financial savings |
| Length | <ul style="list-style-type: none"> ● Reasonable length, and no unnecessary info | <ul style="list-style-type: none"> ● Too long or too much unnecessary information |
| Other | <ul style="list-style-type: none"> ● Use of behavioral insights ● Provision of multiple metrics or useful metrics | <ul style="list-style-type: none"> ● Lack of trustworthiness |

Appendix C. Eye-Tracking Results for Home Energy Assessment Reports (Five-Page Abridged Versions)

REPORT 1: ENVINITY

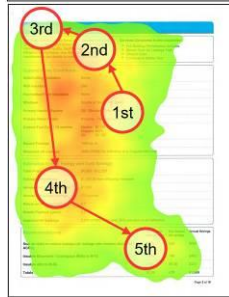


REPORT 2: H.E.A.T. SQUAD



SUMMARY

| | |
|--|--------------------------|
| Building Characteristics | Energy Audit Details |
| Address: 123 Main St, City, State, Zip | Report Number: 12345 |
| Year Built: 2010 | Energy Auditor: John Doe |
| Number of Units: 1 | Report Date: 10/20/2023 |
| Property Type: Single-Family Home | Report Duration: 2 Hours |
| Energy Audit Method: Walk-Through | Report Format: PDF |
| Number of Recommendations: 15 | Report Status: Final |



PRIORITIZED RECOMMENDATIONS

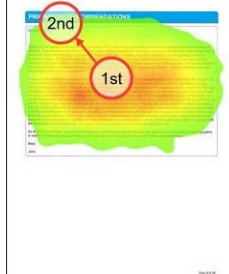
1. Weatherstripping: Seal gaps around doors and windows to prevent air leakage.

2. Insulation: Add insulation to attics and walls to reduce heat loss.

3. Energy-Efficient Lighting: Replace incandescent bulbs with LED bulbs.

4. Programmable Thermostat: Install a programmable thermostat to optimize heating and cooling.

5. Water-Saving Fixtures: Install low-flow showerheads and faucets.



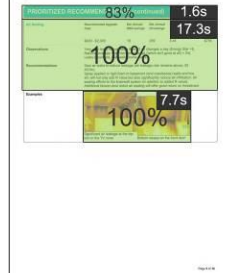
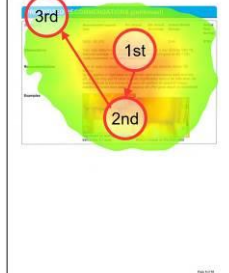
PRIORITIZED RECOMMENDATIONS (continued)

6. Energy-Efficient Appliances: Upgrade to Energy Star rated appliances.

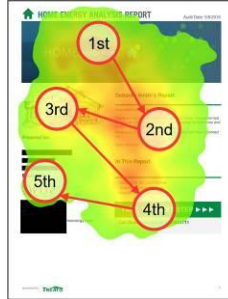
7. Solar Panels: Consider installing solar panels to generate clean energy.

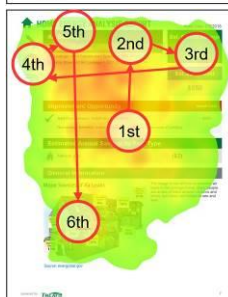
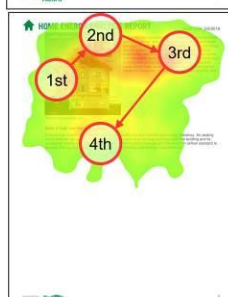
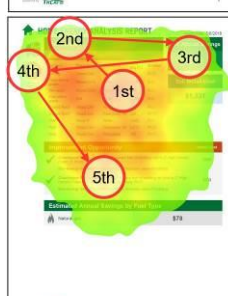
8. Smart Power Strips: Use smart power strips to reduce energy waste from electronics.

9. Energy Audit Follow-Up: Schedule a follow-up audit to track progress.



REPORT 3: WISE HOME ENERGY



Appendix D. Mediation Analysis of Verbal Presentation, Satisfaction with Form of Presentation, and Likelihood of Action

We conducted a mediation analysis to examine whether presenting assessment results verbally (in addition to a written report) has an indirect effect on self-reported likelihood of action, via satisfaction with form of presentation. That is, we sought to determine if the relationship between verbal presentation and likelihood of action was mediated by satisfaction with how results were presented. We depict this relationship in figure D1.

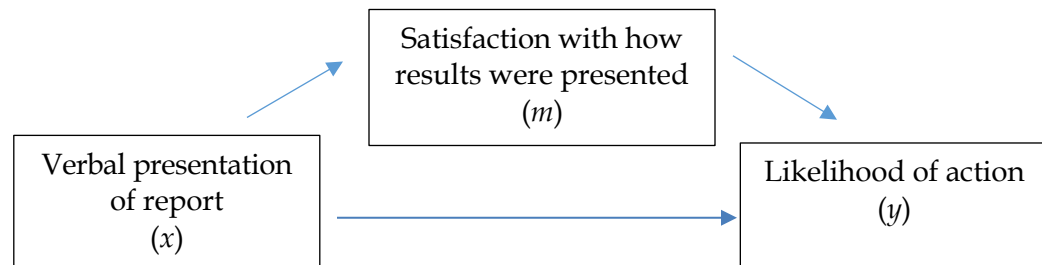


Figure D1. Mediation of relationship between verbal presentations of report (x) on likelihood of action (y) by satisfaction with form of results (m).

Baron and Kenny (1986) describe a “causal steps” approach, suggesting that four criteria are required to establish mediation. Essentially, this involves establishing a statistically significant path between each variable and then changing the strength of one path:

- The initial variable (x , presenting the report both verbally and in written form) must predict the outcome variable (y , likelihood of upgrading the home).
- The initial variable (x) must predict the mediating variable (m , satisfaction with how results were presented).
- The initial variable (x), together with the mediating variable (m), must predict the outcome variable (y).
 - The mediating variable (m) must significantly predict the outcome (y).
 - When the mediating variable (m) is added to the model, the initial variable (x) must become less directly predictive of the outcome variable (y). In most cases, adding the mediating variable to the model causes the relationship between x and y to become nonsignificant.

More recently, however, the causal steps approach has been criticized. In particular, the first requirement of the causal steps approach has been disputed (e.g., Hayes 2009; Shrout and Bolger 2002). Critics suggest that existence of a direct effect of x on y should not be used as a gatekeeper for tests of mediation. That is, one can still conclude that there is an indirect effect of x on y , mediated by m , even if there is no significant direct effect of x on y to begin with. This is the relationship we found in our examination of mediation.

We tested for mediation in our model using a bootstrapping approach (MacKinnon, Lockwood, and Williams 2004) provided by the PROCESS macro created for SPSS by Hayes (2013). Using this approach, we determined:

- x does not directly predict y : $F(1, 440) = .79, p = .37, R^2 < .01; b = 2.42, t(440) = .89, p = .37$
- x predicts m : $F(1, 440) = 20.01, p < .001, R^2 = .04, b = 9.25, t(440) = 4.47, p < .001$
- x and m together predict y : $F(2, 439) = 62.39, p < .001, R^2 = .22$
 - m predicts y : $b = .62, t(104) = 11.12, p < .001$
 - With m in the model, the relationship between x and y , $b = -3.28, t(104) = -1.34, p = .18$, is significantly lower than when m is not in the model ($b = 2.42$ to $b = -3.28$). The direct relationship between x and y is nonsignificant in both cases ($p = .37$ and $p = .18$), but the difference between the paths in each model is significant.

Using the modern criteria for mediation (Hayes 2009), we concluded that indeed there was an indirect effect of presenting assessment results verbally on likelihood of action. This relationship was mediated by satisfaction with how results were presented. Presenting results both verbally and in a written report is associated with higher satisfaction with how results are presented. This, in turn, is associated with a higher self-reported likelihood of action.