A Guide for Field Testing Court Forms & Self-Help Material

Demetrios Karis, Ph.D.

April 30, 2021
(Updated on June 7 and July 13, plus minor updates in August)

demetrios.karis@gmail.com, dkaris@bentley.edu

Executive Summary

There are over 500 legal forms used throughout Massachusetts, and many are difficult to understand and fill out, especially for self-represented litigants with little education. To improve this situation, the Massachusetts Trial Court is in the process of revising forms, starting with forms that are used most frequently.

There is also a variety of self-help printed material, which will also be revised. Many of the forms and self-help material are available online, and the 100 most commonly used forms are currently being converted to HTML5 so they can be filled out online. Some forms can also be submitted online, although this is currently the exception, as even most HTML5 forms need to be printed out or emailed to the court as attachments.

The Chief Justice of the Trial Court and Court Administrator have adopted principles and goals to guide this process. They write that, "Each Trial Court Department shall simplify and standardize forms and self-help materials for external court users with priority given to those forms and self-help materials most commonly used by unrepresented litigants." All forms shall, "include 'plain language' which keeps the intended reader in mind… [and] give clear and simple directions about how to use the form, how to complete the form, what to do with the form, and what to do next."

Although the Trial Court has also developed readability guidelines (see Appendix 2), there is a need to field test all draft forms and self-help materials before they are finalized. In this document, I describe a number of usability techniques that can be used to test forms and documents, whether on paper or online. Based on my experience as a UX (User Experience) practitioner and researcher, as well as the literature on technical communication and usability, I present detailed recommendations about how I think the Trial Court should proceed.

I propose that researchers conduct evaluation sessions either in person sitting next to the participant or observing them over Zoom. Exactly who the researcher and participant should be, and all other details of the procedure, are explained in this document and multiple appendices. There will be slightly different procedures for the different types of material, which include paper forms, HTML5 forms, and both written and video self-help material. The essence of the procedure is simple: create a scenario that requires filling in a form, give a study participant the form, and watch as they attempt to fill it out. Have the participant think out loud while filling in the form. In addition, probe about confusions while the form is filled out, and then interview the participant afterwards. I propose following an iterative process: revise an existing form, test it, identify problems, revise the form, test again, and repeat this process until no new major problems occur.

The evaluation sessions can be carried out either by court personnel, following the recommendations in this document, or by a group of UX volunteers organized and directed by Dr. Demetrios Karis, who has been conducting research for the Trial Court with his Bentley students for the last year and a half. Although the detailed information in this document should be sufficient for setting up and running document usability sessions, I am willing to provide a tutorial on how to use the procedures detailed in this document. Instructions, an informed consent form, an interview guide, and other material that can be used during the evaluation are presented in multiple appendices.

I base the recommendations here on my own usability experience as well as the literature on document usability, which I summarize briefly. These are powerful techniques, and I'm convinced that they will work well. We are now in the process of testing these techniques, and preliminary results confirm that they are effective in uncovering problems in both legal forms and self help material. It has also become clear that we will need to work closely with the Trial Court on this project, in part because we will need to confer with experts in order to create realistic scenarios for our testing. In addition, many parts of the procedure will need to be approved by the court (e.g., the consent form and a demographic questionnaire).

Update, June 7: After experiences over the last month and a half with both usability studies and expert reviews of court forms, it is now clear that there should be some modifications to the recommendations provided here. First, the iterative process I recommend should never start with revising a form; it should always start with either a usability study or an expert review of the form. The results from this work will provide important information about which parts of the form are confusing or need more work, and this information will save time. Second, and as I expected, expert reviews have been incredibly effective in uncovering problems on forms. Given that an expert review can be completed in two or three days, it makes sense to start with an expert review if there is no time for a more formal usability study. You can judge for yourself whether the findings of an expert review are useful by looking at these two expert reviews, which we recently completed:

- [Petition for Appointment of Guardian of Minor: An Expert Review](#)
- [Domestic Relations Summons: An Expert Review](#)

The Petition for Appointment of Guardian of Minor is an existing form, whereas the Domestic Relations Summons is a draft of a revised form.

Table of Contents

Introduction

Litigants must provide information for almost all legal procedures, and they do this by filling out forms provided by the court. Unfortunately, many forms are difficult to understand and complete accurately by self-represented litigants, especially those who have limited English proficiency and limited education. The Massachusetts Trial Court has long recognized this difficulty for litigants, and provides free assistance via lawyers and volunteers in multiple Court Service Centers as well as providing court access and assistance to many organizations that provide free legal services.

Improving the usability of legal forms, plus related self-help material, is perhaps the most efficient method to help litigants and increase equal access to justice, and the Trial Court is now engaged in a massive effort to review, revise, and test the more than 550 legal forms used in Massachusetts. Following plain language guidelines and defining legal terms is essential, but only by testing the forms with representative users can we be sure that they are understandable and easy to use. This document provides detailed procedures for evaluating forms (including online HTML5 forms), self-help material, and explanatory videos. The essence of the procedures involve one-on-one sessions in

which a researcher watches and questions a person as they attempt to complete a form or understand self-help material.

The main purpose of this document is to provide practical advice on how to carry out usability evaluations on court forms and self-help material. Although I ground the techniques based on the research literature, the focus is on the specifics of how to use the techniques in an actual study. In many places, I describe the ideal situation and then present the more practical alternative. For example, researchers and practitioners are in agreement that you should evaluate documentation or a user interface with "representative users" -- people with the essential characteristics of the population most likely to use the system or material (e.g., people with specific prior knowledge, experience, or education). Consider someone trying to get child custody. Their specific prior knowledge will probably include basic information about child custody, and it will certainly include the name and age of the child in question and details about the other parent. Despite the ideal situation of using fully representative users, in many cases a "convenience sample" is a good first step and will provide valuable information. A convenience sample is just a sample of people who are available and easy to use as study participants. They may not have the participant characteristics the researcher is looking for, but it's convenient to use them.

After describing document usability techniques in the main body of the paper, I present examples, procedures, and testing material in multiple appendices. I also provide an annotated bibliography for those who want to delve deeper into these topics, plus appendices on plain language and readability formulas.

An Example of a Confusing Court Form
Difficulty in filling out court forms is a problem in all states and most countries. In England, for example, you can apply for "fee remission" for reduced court fees if you qualify based on income and benefits. But first you must fill out a 2-sided form that comes with 37 pages of instructions! Out of 500,000 applications a year, about 75% get rejected because most people don't understand the legal language and provide incorrect supporting evidence (Santos, 2015). In this case, it took four months of user testing and 11 revisions to improve the form and reduce the 37 pages of instructions to 12 pages, primarily by removing complexity and unnecessary material. For many forms and self-help material, we should be able to complete testing and revision within two or three weeks, although complex forms may take longer.

How to Test Paper Forms

Here is an outline of a procedure I recommend for evaluating forms, followed by sections and appendices that describe the procedure in much more detail. The testing should take place in a quiet room or online over Zoom.

- Welcome, briefing, instructions, informed consent
- Scenario specific to a particular form
- Participant fills out the form while describing what they are thinking, and also marks the form by
  - underlining words, sentences, or instructions to indicate areas of confusion, or writing a ? at the point of confusion
  - writing a plus (+) on the form to indicate a part that is especially clear or helpful
- The researcher probes to understand the participant's difficulties
- After the participant completes the form, the researcher interviews the participant to learn more about the experience and problems with filling out the form
- The researcher debriefs, thanks, and pays the participant
- The form is examined for errors and positive comments, which are documented

The researcher begins a session by welcoming the participant, reading instructions, answering questions about the session, and then having the participant read and sign what's called an "informed consent" form (see Appendix 7). The researcher then explains and demonstrates the procedure that the participant should follow. The researcher next reads a scenario and then the participant fills out a form. For example, if we were testing a form that needs to be filled out to allow parenting time[1], one scenario might explain that the participant should pretend that he has a child but is not married to the mother and has been denied access to the child[2]. We will provide any additional information that a litigant needs or would already know to fill out the form (e.g., one or more addresses).

Then, after these instructions, the participant fills out the form while describing what he is thinking (this is called a "think-aloud protocol"). While filling out the form, the participant is instructed to underline any words, sentences, or instructions that are confusing, or write a ? on the form. He can also write a plus on the form to indicate a section that was especially clear and easy to understand. The researcher should not help the participant while he fills out the form. At the end, or at predetermined points in a

---

[1] Complaint for Custody-Support-Parenting Time Pursuant to G. L. c. 209C; https://www.mass.gov/doc/complaint-for-custody-support-parenting-time-cjd-109/download
[2] For the sake of this example, I am assuming that the participant is male. For female participants we might decide to write a different scenario.

long form, the researcher would also question the participant with pre-determined probes. Here are examples of some probes for the form concerned with parenting time:

- Can you tell me the difference between a Plaintiff and a Defendant?
- "Why did you choose" [say the choice made by the participant]; or if no choice was made, "Why didn't you select one of the choices?"].
- What does it mean to say that a child was "born out of wedlock"?[3]

After the participant fills out the form, marking parts that were confusing and parts that were clear, and answering the probes, the researcher will interview the participant using an interview guide that includes a set of basic questions (see Appendix 8 for an initial draft). The researcher will ask for general impressions about the form, and then ask the participant to point out parts that were particularly confusing. The researcher will also ask whether more general information is needed on the overall legal procedure. Another technique that is effective is to go through the form with the participant after it is completed, reviewing what was straightforward and what was confusing. This typically elicits additional useful information.

After completing the form, the researcher would end the session or continue on to evaluate another form, depending on how long the participant has agreed to participate and the elapsed time. At the end of the session, the researcher thanks and debriefs the participant, explaining what we are doing and why. If there is funding available, the researcher will pay the participant, get a receipt, and say goodbye.

There is a clear criterion of success built into this method: Does the participant fill out the form correctly? In particular, if the participant were to turn the form in to the court, would it be accepted? After the session, a court employee familiar with the form in question could evaluate it and determine whether it would be accepted or rejected, as well as noting where the problems occur. Along with the concept of saturation, described below, the percentage of successfully filled out forms can serve as an indication of whether to continue revising and testing the form or to stop testing and finalize the form.

Testing Forms Remotely
Here is a simple way to test forms remotely that does not require the participant to load or use any software other than Zoom. Both the participant and researcher will have the paper form in front of them. The researcher will also have an electronic version of the

---

[3] Doesn't everyone know what "out of wedlock" means? Maybe, but maybe not; remember, we are not typical self-represented litigants, and what users know and don't know will often be surprising, especially when we consider that many court users have limited English proficiency (LEP).

form on their desktop, and will share the screen with the participant. The researcher can arrange Zoom windows to show both the form and the participant's face.

The participant will see the form on the computer monitor and will fill out the paper version of the form while thinking out loud. In this procedure, the researcher can follow along as the participant fills in the form but can not actually see the participant filling out the form. The researcher can take notes on a copy of the paper form during the session. Although this situation is not ideal, it should work almost as well as other techniques in which the researcher can see the participant filling out the form. After filling out the form, the participant could also take a photo of the form and send it to the researcher. Examining the filled-in form and the marks the participant made while filling it out should help the researcher to understand any confusions or difficulties experienced by the participant.[4] Other techniques for remote testing are discussed below in the section on testing HTML5 forms.

Usability Edits

I am suggesting that we combine the verbal protocol and a mark-up technique described above. This was originally called a user edit, and later a usability edit (Atlas, 1981; Soderston, 1985)[5], and can include one or both of these techniques[6]. The verbal protocol here just means that the participant is instructed to think out loud, describing their thought processes and actions while completing a task. The mark-up technique involves writing pluses and minuses in the margin of a document, or using a variety of simple codes, including the following: circling undefined terms, underlining ambiguous or confusing sentences, crossing out extraneous information, and writing in a star where an example might help. I have simplified this to include just one instruction: underline or circle or write a ? on any parts that are confusing or unclear. The researcher will then question the participant to determine the exact nature of the problem.

Behavioral Protocols

Some researchers use the more general term "protocol-aided revisions" to include both think-aloud protocols such as the usability edit as well as behavioral protocols in which participants do not think aloud but rather complete a task while being observed (Schriver, 1991). Using behavioral protocols, researchers may note, for example, "...where readers look for information (in indexes, in tables of contents, in glossaries); how quickly people can find information (in searching for information online); how users

---

[4] Our initial pilot testing indicates that multiple variations of this technique are effective. In most cases so far, we find that examining the marks a participant makes on the form does not add much to what we learn by listening to participants as they think aloud while filling in the form.

[5] Both Atlas and Soderston describe only the use of a verbal protocol, either alone or in the context of a traditional task-based usability study. Later researchers mention various markup techniques.

[6] Although Atlas is often credited with authoring the first paper on user edits, he points out that his "...'discovery' of the user edit was essentially a re-invention of the wheel" (Atlas, 1998) because the technique was already in use in some companies and the U.S. Army.

make errors and recover from them while operating machinery; how features such as color, windowing, or display rate influence people's ability to use computers; or how quickly and accurately people can perform a task while using a set of instructions (using a manual to assemble a bicycle)" (Schriver, 1991). Behavioral protocols may be useful for court usability work other than with documents, where I think its use is limited. For example, consider trying to understand a litigant's general experience in solving a legal problem such as, "How can I obtain child support?" Behavioral protocols can help in this situation to understand the sequence of actions a person might take. For example, how does a person seek out information and assistance, and in what order? What websites or paper documents are used, are friends or lawyers consulted, and does the individual call or visit a courthouse? Within the UX field, this is often called a customer or user journey map, and is a technique used to help visualize a process (in this case, of obtaining child support).

Problems with a Scenario-based Usability Approach to Legal Forms
The approach described above relies on creating a scenario to match a particular form and then asking the participant to imagine being in the situation described by the scenario (e.g., wanting to request parenting time). For the participant to complete the form, the researcher needs to provide the required information, such as the name and address of the child. The problem is that a great deal of information is required to complete many forms, and providing this information on a sheet of paper constitutes an unnatural situation that may make it much easier to complete the form. In many cases, a major hurdle to completing a form may actually be the step of collecting all of this information, and a form may need to be completed over time as the information is collected.

Consider the situation in which you want to file for guardianship of a minor (https://www.mass.gov/files/documents/2017/01/ud/mpc140-petition-for-appt-of-guardian-for-minor-fill.pdf). The form is filled with requests for names and addresses: of the minor, the biological mother and father, individuals who recently had primary care of the minor, and siblings and relatives. There are also a variety of Yes-No and choice questions (e.g., "Is there a nomination of a guardian by will or other writing signed by a parent or guardian?"). Laying out this information for a study participant in the order it appears is almost like filling in the form for the participant. A better way is to include the information in a long scenario, without using the same terms as in the form (e.g., the scenario would not include the phrase, "nomination of a guardian"). The problem for a participant then becomes understanding the scenario and being able to accurately transfer information from the scenario to the form. How to create scenarios and test these types of forms will need to be resolved via preliminary testing. Even in the unnatural situation of providing a detailed scenario, my prediction is that many problems

will still emerge; e.g., if the participant doesn't understand what "nomination of a guardian by will" means, then this should become clear during the study. However, if this technique doesn't work well for some forms, then it may be necessary to collect more naturalistic information by studying errors in actual forms turned in to the court, as well as other procedures, as described below in the section on Initial Form Revision.

Expertise Required to Create Scenarios
The other UX Researchers and I can create scenarios for each form, but we will need the expertise of court employees to finalize them. We will need to understand the most typical context and background of a person filling out a particular form, whether a form needs to be submitted along with another form or an affidavit, and whether the form is one of a sequence of forms that needs to be filled out during the course of a legal procedure. We will also need to know what information is required and what is optional, as this is sometimes unclear from the forms themselves. We don't want to provide a participant with all the optional information, because we want to learn during the evaluation whether a person understands what is mandatory and what is optional. If we provide optional information this will be difficult to discern.

An Alternative: More General Usability Testing
Filling out the form *Petition for Appointment of Guardian of Minor* is just one step in the process of becoming a guardian. There may be multiple forms required (see https://www.mass.gov/lists/court-forms-for-guardianship-of-a-minor), and the process for an individual seeking guardianship probably starts with gathering information and trying to understand the process. The process requires many steps and unfolds over time. An alternative to just testing a single form, which is what I am recommending in this document, is to conduct a more general session (or multiple sessions) on guardianship. The participant in such a study would also start with a scenario, but would not be given a form. The participant would need to go through mass.gov to learn about the process, decide what forms to fill out, and then try to fill them out. This would take much more time, but the advantage is that it would provide information about multiple forms as well as the related explanatory and self-help material on mass.gov.

Testing Legal Terms
It may be efficient to run a quick study on legal terms found in court documents before actually testing forms or self-help material. A fair number of people will probably be unclear on the distinction between "Plaintiff" and "Defendant," but practically no one will be able to define "affiant," which is found on some forms. There are multiple ways to run such a test. The easiest is just to ask a participant to define a list of terms. A more accurate way is to give participants a form and ask them to define a term used in the form. For example, ask participants to define *affiant* after giving them the form Affidavit

of Address Re-verification Upon Request for Default Judgment (https://www.mass.gov/doc/affidavit-of-address-re-verification-upon-request-for-default-judgment/download), which starts like this: "The affiant in this action states that the defendant's residential address was re-verified by within three months prior to the request for default judgment by at least one of the following three methods, as indicated below." We could also include non-legal but difficult vocabulary in this test (e.g., consanguinity).

How to Test HTML5 and Fillable PDF Forms

The procedure for testing online HTML5 and fillable PDF forms is almost identical to the procedure described above for testing paper forms. In the ideal situation, the researcher would observe the participant in their own living space or office, but since that is typically difficult and time consuming for the researcher, most testing will be done either remotely, or in a courthouse with the participant using a court computer.

During a Zoom call, we will ask participants to fill in forms after sharing their screen while thinking out loud. In this case, the researcher can see the form and what the participant is doing while listening to what they are saying. The participant can not, however, underline confusing sections or write pluses on the form).[7]

The participant would be asked to complete the form, print it out, and also email it to the researcher (because printing out and emailing a form are common tasks litigants must perform). The researcher or a court employee can judge whether the participant has succeeded or failed in accurately completing the form and then printing it out and emailing it.

How to Test Written Self-Help Material

Here is an outline of a procedure to test written self-help material, followed by sections that provide specific details. There is significant overlap with the procedure outlined above for forms, as both procedures involve many common parts of an evaluation session.

---

[7] If HTML5 or fillable PDF forms are not available, then the participant will need to use an application to edit the form for online testing (or use paper as described in the section above). If the participant has Word or some other word processing application this should be easy (we would create a Word version of the form), and as with the HTML5 forms, the participant would share their screen while filling out the form. Many people will not have Word, however, and in that case there are other technical approaches we could consider.

- Welcome, briefing, instructions, informed consent
- Scenario specific to the self-help material being tested
- If the material is longer than a paragraph, it is divided into sections by the researcher before the session begins
- Participant reads the first section out loud and also marks it by
  - underlining words, sentences, or paragraphs, and adding question marks to indicate areas of confusion
  - writing a plus (+) on the form to indicate a part that is especially clear or helpful
- The researcher asks the participant to paraphrase the section in their own words
- The participant then continues to the second section and the process repeats
- The researcher may also probe to understand whether the participant understands key terms or concepts
- After the participant completes going through all sections of the self-help material, the researcher interviews the participant to learn more about the experience of reading and understanding the material
- The researcher debriefs, thanks, and pays the participant

Although the participant will not be instructed to think out loud, if he or she makes any comments while reading a section the researcher will encourage more comments. Just observing the participant read the material may be informative, based on pauses, reading speed, facial expressions, and whether the participant chooses to reread a section. Imagine, for example, a participant reading a sentence, stopping, scowling, and then rereading it. This is an indication that something is wrong, and the researcher can probe to discover the source of the problem.

Appendix 11 provides an example of how to test the FAQ about the Servicemembers cases in the Land Court. My prediction is that the first sentence, given its length and qualifying clauses, will be confusing, and this will be obvious by observing the participant's reading behavior and facial expressions.

Combining the Evaluation of Forms and Self-Help Material
Some self-help material is associated with a particular form, and it may make sense to test both simultaneously. For example, if you want to have court fees waived you need to file an *Affidavit of Indigency* (https://www.mass.gov/doc/affidavit-of-indigency/download). There is a separate document with *Instructions for Users of Affidavit of Indigency and it's Supplement* (https://www.mass.gov/doc/instructions-for-users-of-affidavit-of-indigency-and-its-supplement/download). Rather than test each individually, a participant could be given both,

which would provide a more realistic situation. Many individuals might be unlikely to read through the entire instructions, but what they do read would be instructive. To obtain information on the entire instruction document, the participant could be instructed to return to it after attempting to complete the form and read through the other sections.

How to Test Video Self-Help Material

As with the other types of material, I suggest testing the videos in one-on-one sessions. In this case, as with self-help material, we would not use a think-aloud protocol. The videos tend to be short, so it makes sense to let a participant watch the entire video and then ask them to summarize what they learned. We will probably instruct participants to watch the video in whatever way is normal for them. That is, they will be in charge of the video controls and can pause or replay if they want.

For each video, we would create a list of main points ahead of time and then compare the participant's summary to our list of main points. We can also create questions based on the main points if they are not included in the participant's summary (e.g., Can you tell me whether it's possible to appeal a ruling in small claims court?).

Consider the small claims videos (https://www.mass.gov/lists/small-claims-videos). Most of these videos are quite short, two minutes or less, while a few are up to four minutes long. After the participant presents a summary, the researcher would interview the participant, asking for more detail about the summary to make sure the participant understood the main points. The interview could also include questions about parts of the video that were unclear, as well as collecting the participant's feelings and attitudes about a number of attributes of the video: the speaker, whether he or she was speaking intelligibly, the pace of the video, the amount of information presented, the quality of the text and graphics, and so on. Before starting to evaluate videos we should create a standard list of video attributes to cover.

How to Test Translated Forms & Self-Help Material

After a form or self-help material is revised and tested, it may be translated into multiple languages. Translated forms and material should be tested and revised in the same way as the original form. When a revised form in English with no serious problems is translated into another language, new problems may be introduced, and the best way to identify these problems is by using the same techniques as described above. The participants, of course, should be native speakers of the new language. Some participants should also have no more than a highschool education and come from low-income households.

The Benefits of an Expert Review

Before starting user testing, many problems can be identified via an expert review (see https://www.nngroup.com/articles/ux-expert-reviews/ for a description of expert reviews). In my experience, which is backed up by research, expert reviews are often as effective as usability tests in identifying problems. They are also much quicker and cheaper, as they involve having one or more experts examine a product or service (or form in this case) to identify problems. The experts do this based on a combination of their own experience as well as user interface and design principles and heuristics.

To illustrate this, consider the HTML5 form for Child Support Guidelines Worksheet (https://courtforms.jud.state.ma.us/publicforms/PFC0001). In five minutes, I found what may turn out to be several problems. There's an option to *Reset Form*, which makes sense. But in other places on the form, if you try to correct a mistake you've made, the form is reset! For example, if a person fills in the top of the form, then for the second question checks the first radio button, but reconsiders and realizes that the second radio button should have been checked, the form will be reset and the person will have to re-enter all the information at the top! In other forms, if the radio buttons are farther down, even more work could be lost.[8] There is also an instruction that, "You must fill in the blanks in the form in order, or the form will not work." It's not clear what it means that "the form will not work." In what way won't it work? And what does it mean to fill the fields in order? In some situations it may not be clear whether to first go down and then over, or whether you should proceed left to right and then down. In addition, there are instructions at the top, but they are general instructions for all HTML5 forms, not this particular form, and they could be improved. We may also want to test a form's compatibility with screen readers such as NVDA, JAWS, VoiceOver, or Narrator, as well as screen magnifiers.

Consider another example. If you want to annul your marriage you need to fill out a Complaint for Annulment. After entering in the address of the Plaintiff and the Defendant, the next section asks you to, "Please check and complete ONLY ONE of the following sections. [The numbers below are checkboxes on the form.]

1. On _____ (date) the parties went through a marriage ceremony at _____ (City/Town), Massachusetts

---

[8] I realize that changing options that indicate who has financial responsibility for the children may have implications with respect to the calculations further down the form and that some fields may need to be reset, but it seems unlikely that basic information at the top of the form would also need to also be reset.

2. On _____ (date) the parties went through a marriage ceremony at _____ (City/Town), _____ (State), at which time the plaintiff was domiciled in the Commonwealth of Massachusetts and he/she is domiciled in the Commonwealth of Massachusetts at the commencement of this action.

3. On _____ (date) the parties went through a marriage ceremony at _____ (City/Town), _____ (State), and the plaintiff has resided in the Commonwealth of Massachusetts for five (5) years immediately preceding the commencement of this action.

These three options do not seem to be logically distinct. Can "Massachusetts" be filled in as the State in the second or third option? Can't the third option be a subset of the second? When testing this form, neither the participants nor I had any idea about how to explain the distinctions between the three options. Upon further reflection, I think this is the intent: The second option is for a marriage in another state, and while at the time of the marriage and when this request for an annulment was made the plaintiff lived in MA. The third option covers the situation in which the marriage was in another state and took place while the plaintiff did not live in Massachusetts, but he or she has now lived in Massachusetts for five years. I'm not sure this interpretation is correct, but it makes sense, although it is impossible to determine definitively from the description of the three options. In extreme cases like this, it will be clear to any UX researcher (and probably to anyone!) that the options are confusing and should be rewritten. No user testing is required.

Interview-Based Approaches to Form Submission

There are at least two online systems that walk a litigant through a legal process, asking questions, providing definitions and explanations, and then filling out and submitting the form to the court at the end. One is by the Suffolk Law's Legal Innovation and Technology Lab, while the other is Odyssey Guide & File by Tyler Technologies. Some pages on mass.gov link to these systems. For example, from https://www.mass.gov/how-to/file-a-small-claim-in-the-boston-municipal-court-district-court-or-housing-court: "The easiest way to file a small claim is to file online using guide and file (https://massachusetts.tylerhost.net/SRL/SRL/ExecuteInterview). The program will walk you through how to fill out your forms using plain language."

There are also links from mass.gov to the interviews created by the Legal Innovation and Technology Lab (https://courtformsonline.org/).

These systems should also be evaluated by running usability tests, but that is outside of the scope of this project.

Metrics: Measuring Form Usability and Problems

There are many ways to measure how easy it is to accurately fill out a form and to track the number of errors. "Summative" usability testing involves studying a relatively large number of participants as they try to fill out forms and then keep track of success rate, the number of errors, and the time required to complete the form. A relatively large number of participants is required in order to obtain an accurate estimate of success rates and other measures. Given the time required, I suggest considering other methods, including the following:

- Analyze all copies (or a sample) of a particular form when they are turned in. How many errors are there on the forms, and what percentage of forms are rejected due to serious errors or omissions?
- Examine the number of times a particular form is downloaded versus the number that are turned into the court. Consider an example: let's assume that five forms are printed for each one that is turned into the court, while after the form is revised only two are printed for each that is turned in. The court would save printing costs, and one explanation is that with more difficult forms people make mistakes and have to start over with a new form. Evidence for exactly why there is a change in the number of downloads versus turned-in forms could be obtained via usability studies, interviews, and other means.
- Keep track of questions and requests for help for a particular form at the Court Service Centers, the Virtual Registry of the Probate and Family Court, and at clerk's offices for other departments. A sampling system could be set up such that only a sample of questions and requests for help are tracked. The same type of tracking system could also be set up as part of the Lawyer for the Day program and various legal aid organizations
- [I'll add other options here]

Some of these methods are ideal for a court intern or for research collaborations with local colleges and universities.

Usability testing, as defined in this document (even with a small sample size), is the best way to determine whether a revised form is an improvement over the existing form. However, it would be very useful and informative to augment the results from usability testing by tracking performance on the measures above both before and after a form is revised.

General Methodology

Many aspects of the evaluation methodology will be identical across different materials and procedures, and in this section I discuss who the participants should be, how many to include in each evaluation, how to recruit them, and the iterative design process. I start by presenting the advantages of creating a problem database and initial ideas for a problem taxonomy.

A Problem Database

In trying to fill out a form, or learn from self-help material, a person may experience many different types of problems, and it may be useful not just to categorize these problems, but to create a database of problems and solutions that can be applied to forms that have not yet been tested.

After completing the evaluation of each form, we should examine the problems and confusions encountered and make a judgment about whether they might also cause problems in other forms. Here is a simple example: many forms use the terms *Plaintiff* and *Defendant*. If we learn early on that people don't understand the difference and an explanation is required, then it is likely that this will be a problem in many other forms that we have not yet evaluated. Before starting a new evaluation, we would go through our problem database to make sure that previous problems would not arise again. If this technique is effective, then evaluating forms later in the process should find fewer problems, require fewer participants, and be completed more quickly. We will, in effect, be creating a new set of very specific guidelines that complement the Trial Court Readability Guidelines (presented in Appendix 2).

A Problem Taxonomy

A problem taxonomy is just a classification of the different types of problems users encounter when trying to fill out forms or understand self-help material. Here are some ideas for a problem taxonomy, which can also be used to create a coding scheme.

- **Comprehension**: People may not understand the material due to the syntax or vocabulary used.
    - A subcategory is the use of technical terms that are not understood
    - Comprehension problems may also arise from ambiguous information
- **Missing Information/Completeness**: People may need additional information to understand the material
- **Misleading instructions**: Sometimes it is difficult to write verbal instructions without flowcharts, figures, or diagrams
- **Incorrect information**
- **Key information may not be visible**

- **Procedural problems**: For example, missing steps in a procedure or too many steps in the procedure
- **Relevance**: More information is included than is required, or extraneous information is included
- **Problems with examples**: examples may be difficult to follow, not be useful, or contain inaccuracies or elements that are not relevant
- **Structure and formatting**: People may have problems based on the structure or formatting of the information, including the order in which information is presented. Typography and section headings may also lead to problems.
- **Graphic Design**: The layout or illustrations may be confusing; the use of tables, charts, and other visual aids may be difficult to understand
- **Workflow** in paper and online forms
- **Accessibility**

Iterative Design & "Saturation"

I suggest that we follow an iterative design process: as we collect information on a form, we will update it when necessary, even after just a single participant if there is a clear flaw. We will then test again, update the form again, and continue this process until we are satisfied that we have eliminated all serious problems[9] (note, however, the alternative described in the section below). The concept of "saturation" arises here. Saturation is an important concept in determining how many people to include in a usability study, and this concept is also important in determining how many participants to interview. When the researcher learns nothing new and starts hearing the same problems (in usability testing) or the same themes (in interviewing), then you have reached saturation. Exactly how to define saturation is sometimes an issue, and often discussed in the literature. (For information on saturation, see the classic article by Guest, Bunce, & Johnson, 2006; for a discussion of saturation and related points, see Baker & Edwards, 2012.)

An Iterative Design Process[10]

Here is my outline of an iterative design process that we can follow. (This process can also be used for testing self-help material.):
1. Before testing with users, conduct an expert review of the form (or set of forms in one area, such as guardianship)

---

[9] This is often referred to as the RITE Method, or Rapid Iterative Testing and Evaluation.

[10] Schriver (1991) provides a more detailed process model of what she calls a protocol-aided revision. The model she describes "...is intended to help writers/document designers see the relationship among three components: (1) cognitive processes in protocol-aided revision, (2) writer/document designer activities while engaged in these processes, and (3) outputs of processes and activities." It may be worth discussing this model among the different individuals and groups involved in this project.

2. Check the elements of the form against the Problem Database to determine whether there are any other likely problems that should be fixed
3. Revise the form based on the expert review and the results from checking against the Problem Database, as well as the Trial Court's Readability Guidelines, Plain Language guidelines, and any other best practices that are available
4. Test the revised form with representative users using the verbal protocol and markup technique
5. Revise the form based on the results of the testing (i.e., fix the problems identified)
6. Test again with new participants
7. Revise the form again based on the results of the testing
8. Repeat the process of testing and revising the form until no new problems arise from testing
9. After completing the process for a particular form, consolidate problems across participants for that form and add them to the Problem Database.

Steps 1 and 2 above should probably be carried out simultaneously. The revision in step 5 may occur after each participant or after two or more participants, depending on the schedule of testing participants and the number and seriousness of the problems uncovered. That is, if after testing the first participant you discover a serious problem that you're sure will need to be fixed, then fix it immediately before further testing.

Please note that this process is not as complex as it may seem, and can be reduced to these five essential steps:
1. Examine the original form for problems based on an expert review, plain language guidelines, and best practices established by the Trial Court.
2. Revise the form accordingly
3. Test the revised form with representative users
4. Revise the form based on the results of testing
5. Repeat the process of testing and then revising the form until no new serious problems emerge (or you run out of time!)

An Alternative: Although an iterative design process with multiple stages is often ideal in situations such as this, it requires close interaction and cooperation between the design (writing in this case) and testing groups. When this is not possible, then testing will continue until saturation is reached and all the problems and recommendations will be delivered at once.

The process of revising and testing hundreds of forms will be a long one. Many forms will have similar problems (e.g., undefined legal terms, a confusing choice of options), and by creating a problem database and fixing problems on forms not yet tested the process will be greatly accelerated.

Initial Form Revision

The court team working on the initial form revisions will include personnel with expertise in both documentation writing and the legal practices and procedures of each court department. Ideally, each court department should collect information on problems and confusions with the forms and self-help material before user testing begins. There are several options (and I assume that some of these approaches may already be occurring):

- Collecting and analyzing forms that have been turned into the court but rejected because they have been filled out incorrectly or information is missing. Understanding what parts of the form lead to the main problems will help in revising the forms.
- Interviewing lawyers working in the court service centers to learn what parts of each form are the most problematic. An even better alternative, although more time consuming, is to instruct all court service center lawyers and volunteers to take notes after assisting litigants on any problems or suggestions for improvements to a form.
- Information on difficulties filling out forms can also be collected from the Lawyer for the Day program and nonprofits that work with low-income litigants (e.g., Rosie's Place has several types of legal clinics)

Participants

The target population is low-income, self-represented litigants who do not have more than a high-school education. Ideally, we would draw all participants from this population. If members of this population can use the forms and self-help material easily, then the probability is very high that others will also be able to use them easily. Many problems, however, will be found by almost any possible participant, even many who work for the court or have a legal background. My suggestion is thus to start testing with a convenience sample of whomever is available, including friends, colleagues, and family members of the researchers and the court employees involved in this project. This will save both time and money. I suggest we start by testing a form with two or three participants from a convenience sample and then continue with two or three participants from the target population. After testing these people, researchers will make a decision about whether to continue testing or to stop.

July Update: We have confirmed that there is no need to recruit representative users for the first round of testing. We have been very effective in uncovering a variety of problems on every form we tested even when using a convenience sample of highly educated participants. Some of these participants even had extensive knowledge of the legal system.

The Ideal Situation and More Information on Using Convenience Samples: The ideal situation is to recruit participants who actually have a need to complete a form or use self-help material. We could do this by recruiting people at a Court Service Center who are trying to get help with a form that we are testing. Before providing assistance, we could try to recruit the person to be in our study. This is ideal because the participant does not have to listen to the scenario we present and pretend to be in that situation while completing the form. The person is in the exact situation we are simulating and has the knowledge and information -- and emotion -- of a person in need of obtaining the particular legal assistance or outcome that the form provides. This is ideal, but I am not suggesting we do this, as we can find the most serious problems with much easier procedures. In fact, although we should always include some participants who meet the characteristics of our target population, in my experience even non-representative users are almost always a good way to start.[11]

The usefulness of a convenience sample demonstrates the power of document usability via a think-aloud protocol, and many others also note the advantages of using a convenience sample. Atlas (1981), for example, writes that, "...even though your best bet is to test *typical* users...you can still get good results by testing users who are much less typical (like people who work for you or your company)." In fact, Atlas (1988) wrote a retrospective on the user edit technique in which he recounts his first experience with the user edit during a pilot test. For that test, they used the document writer's assistant as a participant. She was very familiar with the manual they were testing, and as a result, Atlas and the document's author did not expect her to have any trouble using the manual. However, they "...were wrong. Sentences that seemed to make perfect sense led to confusion, the table of contents led to dead ends, and the information the test subject needed 'hid in plain sight' under headings that were accurate but useless" (p. 22).

How Many Participants?
How many participants should attempt to complete each form? Given the iterative design procedure we will be following, it is not possible to agree on a fixed number

---

[11] I first became aware of the usefulness of a convenience sample in a usability study of GTE Airfone, the first public telephones in commercial aircraft. After spending a lot of time and money to recruit executives who flew frequently (the target population), we found that with a convenience sample of random adults or students from a local university we found the exact same set of problems.

ahead of time. Based on my previous experience, we may only need to test with as few as five or six participants for each form, at least initially. This will depend on reaching saturation, as described above, plus on how much we have already learned from evaluating previous forms. Using a problem database, we may require fewer participants for forms we evaluate later in the process. My experience is in accord with that of Schriver (1991), who reports that document designers at the Carnegie Mellon Communication Design Center typically used about five participants per design cycle.[12]

Recruiting Participants

Low income, self-represented litigants should be able to understand and fill out legal forms without assistance. Some of these litigants will be immigrants with limited English proficiency, and they too should be able to complete these forms successfully. For that reason, we should recruit study participants from these groups. We can recruit participants in multiple ways; for example, from people entering a courthouse, from notices included in letters sent to litigants, from the Court Service Centers, from social service agencies that work with self-represented litigants, from social media, and so on. If we have funding to pay participants, recruitment will be easy. Without funding, recruitment will still be possible but will be more difficult.

Researchers

To evaluate multiple forms and self-help material we will need many researchers. I can train court personnel to run the evaluation procedure described in this document. I have a team of volunteer researchers who can also run these procedures. My team is composed of Bentley UX students, former students, as well as experienced UX researchers who are willing to volunteer their time for this project.

Testing on Desktops/Laptops, Tablets, and Smartphones

Most forms and self-help material is available online, and a person can fill out HTML5 forms online. Some forms can even be submitted online. We know that many litigants will use tablet computers and smartphones for accessing and filling out online forms and other material, so it will be imperative to test on these devices. Unless care is taken during development, problems may occur when using smartphones and, to a lesser extent, tablets. Although testing on different devices (and operating systems and browsers) is important, I see this as secondary to the primary focus on the usability of the forms themselves.

---

[12] Schriver reports that their experience is that approximately half the problems are discovered in each revision cycle, and typically two or three cycles are required.

Steps in Running a Document Usability Evaluation

All studies will involve the following steps. Some are explained below and more information is provided in several appendices:

- Greeting
- Briefing
- Informed consent
- Instructions
- Testing of form or self-help material
- Participant interview
- Short demographic questionnaire
- Debriefing, payment, and thanks

Informed Consent

An informed consent form is presented in Appendix 7. A consent form provides basic information about the study, discusses confidentiality, lists any risks associated with the study, describes remuneration if there is any, and informs the participant that he or she can withdraw at any time. It is often adequate to obtain oral consent during a remote testing session, although sometimes a written signature is required. Traditionally, researchers used the postal service to mail a consent form to participants in remote studies, but today electronic techniques are more common. To obtain a signature in remote studies there is a 30 day free trial available for DocuSign, but I would prefer a more permanent solution. There is also a signature option in Qualtrics that we could use. We can also email the consent form and have the participant print it out, sign it, scan or photograph it, and email it back, although this requires some degree of technical sophistication.

Moderating Guidelines

We will follow moderating guidelines common in usability studies (see, for example, Dumas & Loring, 2008). The key is not to help the participant, not to talk too much, and to remind the participant to think out loud if they remain silent for too long. There are now many excellent books on usability (Rubin and Chisnell, 2008, is a good all-round book).

Interview Guide

An interview guide is presented in Appendix 8. It contains questions to guide the discussion after a participant has filled out a form or read through some self-help material.

Demographic Questionnaire

It will be useful to know demographic information about study participants to make sure we are evaluating the forms with a wide cross-section of the population. The court has already developed a few demographic questions for their Access and Fairness Survey, and we should use these questions in addition to developing new ones. It would be ideal to collect a complete set of demographic information, including age, sex and gender, ethnicity, marital status, location, education, employment status, income, and primary language. It might also be useful to collect information on technical expertise and computer availability (i.e., does the respondent own a desktop or laptop computer, a tablet, and/or a smartphone, and is broadband Internet available from home).

References

Atlas, M. A. (1981). The User Edit: Making Manuals Easier to Use.*IEEE Transactions on Professional Communication*, *24* (1), 28-29.

Atlas, M. A. (1998). The User Edit Revisited, or "If we're So Smart, Why Ain't We Rich?". *Journal of Computer documentation, 22* (3), 21-24.

Baker, S. E., & Edwards, R. (2012). How many qualitative interviews is enough? Expert voices and early career reflections on sampling and cases in qualitative research. *National Centre for Research Methods Review Paper* http://eprints.ncrm.ac.uk/2273/4/how_many_interviews.pdf

Casper, R. A., Lessler, J. T., & Willis, G. B. (1999). Cognitive Interviewing: A "How to" Guide. Reducing Survey Error through Research on the Cognitive and Decision Processes in Surveys. *Short course presented at the 1999 Meeting of the American Statistical Association.* http://www.chime.ucla.edu/publications/docs/cognitive%20interviewing%20guide.pdf

Charrow, R. P., & Charrow, V. R. (1979). Making Legal Language Understandable: A Psycholinguistic Study of Jury Instructions. *Columbia Law Review*, *79* (7), 1306-1374.

Crossley, S. A., Skalicky, S., and Dascalu, M. (2019). Moving beyond classic readability formulas: new methods and new models. *Journal of Research in Reading*, *42*, 541–561. https://doi.org/10.1111/1467-9817.12283.

de Jong, M. & Schellens, P. J. (2000). Toward a document evaluation methodology: what does research tell us about the validity and reliability of evaluation methods?, *IEEE Transactions on Professional Communication*, *43* (3), 242-260, doi: 10.1109/47.867941.

Dumas, J., & Loring, B. (2008). *Moderating Usability Tests: Principles & Practices for Interacting*. Morgan Kaufmann Publishers.

Guest, G., Bunce, A., & Johnson, L. (2006). How many interviews are enough? An experiment with data saturation and variability. *Field Methods*, *18* (1), 59-82.

Hayes, J. R. (1998). Atlas's "The User Edit": The Impact on Product Assessment. *Journal of Computer Documentation, 22* (3), 7-9.

Jarrett, C, & Gaffney, G. (2008). *Forms that Work: Designing Web Forms for Usability*. Morgan Kaufmann.

Jarrett, C., & Reddish, J. (2019). Readability Formulas: 7 Reasons to Avoid Them and What to Do Instead. *UXmatters*, https://www.uxmatters.com/mt/archives/2019/07/readability-formulas-7-reasons-to-avoid-them-and-what-to-do-instead.php

Jarrett, C., & Reddish, J. (2020). How to Test the Usability of Documents. *UX Matters*, https://www.uxmatters.com/mt/archives/2020/05/how-to-test-the-usability-of-documents.php

Lewis, J. R. (2006). Effectiveness of Various Automated Readability Measures for the Competitive Evaluation of User Documentation. *Proceedings of the Human Factors and Ergonomics Society*, 624-628.

Reddish, J. (2010). Technical Communication and Usability: Intertwined Strands and Mutual Influences. *IEEE Transactions on Professional Communication, 53* (3), 191-201.

Reddish, J. (2012). *Letting Go of the Words: Writing Web Content that Works*. Morgan Kaufmann Publishers, Second Edition

Rubin, J., & Chisnell, D. (2008). *Handbook of Usability Testing: How to Plan, Design, and Conduct Effective Tests* (Second Edition). Wiley Publishing, Inc.

Santos, A. C. (2015). Helping People with Court Fees. https://mojdigital.blog.gov.uk/2015/07/29/helping-people-with-court-fees/

Schriver, K. (1991). Plain language for expert or lay audiences: Designing text using protocol-aided revision. In Steinberg, E. R., *Plain Language: Principles and Practice*, pp. 148-172. https://www.researchgate.net/publication/285770693

Schriver, K. (1998). Reflecting on Atlas' User Edit: Changes in Thinking about Usability Between 1981 and 1998. *Journal of Computer Documentation, 22* (3), 10-14.

Soderston, C. (1985). The usability edit: A new level. *Technical Communication, 32* (1), pp. 16-18.

Wishbow, N. (1998). Still Looking for Troubles: Commentary on Marshall Atlas's "The User Edit". *Journal of Computer Documentation, 22* (3), 15-20.

Wroblewski, L. (2008). *Web Form Design: Filling in the Blanks*. Rosenfeld Media.

Appendix 1: An Annotated Bibliography

I highly recommend the articles below that have the author and date in bold black font. The full citations are in the Reference list above.

Atlas (1981). *The User Edit: Making Manuals Easier to Use*
This article is clear and concise (it is only two pages!) and is one of the first to describe the user edit technique for document usability. It is still relevant and highly recommended. In 1998, in the *Journal of Computer Documentation*, Atlas presented "A retrospective look at the discovery, explanation, and fate of user edits." There were also short (three to four pages) commentaries by Hayes (1998), Schriver (1998), and Wishbow (1998) in the same issue. Atlas' retrospective is worth reading. The three commentaries, which are listed in the reference list, are not as interesting.

Crossley, S. A., Skalicky, S., and Dascalu, M. (2019). *Moving beyond classic readability formulas: new methods and new models.*
This article provides a short but comprehensive review of text readability and the problems of classic readability formulas. The authors then develop new readability formulas based on advanced natural language processing tools. See the paragraph in Appendix 14 for more information.

de Jong, M. & Schellens, P. J. (2000). *Toward a document evaluation methodology: what does research tell us about the validity and reliability of evaluation methods?*
The authors discuss the validity and reliability of multiple evaluation methods, but this article is of little practical use. The authors mention and describe 21 methods, although only a few are potential candidates for evaluating court forms and self-help materials. The authors describe research on evaluating questionnaires, which is relevant, but also discuss the results from many studies that focus on the methodologies for evaluating user interfaces, and many of these techniques are not applicable to document usability. A major problem in the paper is its failure to consider the ways in which most evaluation is done in practice; that is, in the context of an iterative design cycle. The main reason for including this article here is the review of 76 research papers, although they only cover the period up until the publication date of 2000.

Jarrett, C, & Gaffney, G. (2008). *Forms that Work: Designing Web Forms for Usability*. Morgan Kaufmann.
Although this book is about web forms, many of the guidelines and best practices, as well as the examples, also apply to paper forms.

Jarrett, C., & Reddish, J. (2019). *Readability Formulas: 7 Reasons to Avoid Them and What to Do Instead*

This is a column on UXmatters that clearly explains what readability formulas are and why you should avoid them. Highly recommended.

Jarrett, C., & Reddish, J. (2020). *How to Test the Usability of Documents.*
This is another column on UXmatters that explains and provides examples of the three main ways to test the usability of documents. Highly recommended.

Reddish (2010). *Technical Communication and Usability: Intertwined Strands and Mutual Influences*
This relatively short (nine page) article by Redish provides an excellent history on evaluating the usability of documents and the intertwined history of technical communicators and usability specialists. Highly recommended.

Schriver, K (1991). *Plain language for expert or lay audiences: Designing text using protocol-aided revision*
This is an extensive (40 page) review of how to test and revise documents, primarily focusing on think-aloud protocols. The article also includes two case studies, including the verbal protocols obtained while reading instructions or general information. The examples are interesting and worth reading, while the rest of the article is not as important.

Wroblewski, L. (2008). *Web Form Design: Filling in the Blanks*. Rosenfeld Media.
Although this book is about web forms, many of the guidelines and best practices, as well as the examples, also apply to paper forms.

Appendix 2: Massachusetts Trial Court Guidelines, Principles, & Goals

There are 34 guidelines in the *Massachusetts Trial Court Readability Guidelines for Printed Self-Help Materials and Forms*:

MA Court Readability Guidelines

The Chief Justice of the Trial Court and Court Administrator adopted the principles and goals listed in the following one-page document: *Massachusetts Trial Court Principles and Goals for Revising and Creating Forms and Self-Help Materials for External Court Users*: Principles and Goals for Revising Forms and Self-Help Material

The guidelines and principles are excellent, although I think the first guideline, on using readability formulas to calculate grade level, should be removed. See Appendix 13 for an explanation.

Appendix 3: Examples of Paper (static) Forms

Here is a sample of the most widely used static forms, based on the number of downloads from mass.gov: [I need to replace these with URLs that are easier to open; this is a major usability problem that the court is aware of and is slowly fixing; see https://www.mass.gov/service-details/what-to-do-if-you-cant-open-court-pdfs. The first URL below should open without a problem. From Chrome on a PC I can open the last three documents by downloading the file and then right clicking and selecting Open with Adobe Acrobat Reader.]

- 2018 Child Support Guidelines: https://www.mass.gov/info-details/probate-and-family-court-2018-child-support-guidelines-worksheet-cjd-304
- Financial statement short form: https://www.mass.gov/doc/courts-dwnld-financial-statement-short-form-cjd-301s/download
- Request for copies form: https://www.mass.gov/doc/courts-dwnld-request-for-copies/download
- Financial statement long form: https://www.mass.gov/doc/courts-dwnld-financial-statement-long-form-cjd-301l/download

Appendix 4: Examples of HTML5 Forms

Here is a sample of HTML5 forms. The first three forms below are all Probate and Family Court forms. They are three of the most used forms across the court system. The fourth form is used across multiple departments.

.
- ● Child Support Guidelines Worksheet:
  https://courtforms.jud.state.ma.us/publicforms/PFC0001
- ● Financial Statement Short Form:
  https://courtforms.jud.state.ma.us/publicforms/PFC0002
- ● Request for Copies (Probate and Family Court):
  https://courtforms.jud.state.ma.us/publicforms/PFC0003
- ● Motion to Vacate Human Trafficking:
  https://courtforms.jud.state.ma.us/publicforms/TC0009

Here are the current instructions at the top of these forms. There are no bullets or spacing to separate the four main points. The formatting should be improved, and the content evaluated as part of user testing.

> "Kindly fill the form and click the 'Save as PDF' Button.The 'RESET' Button clears all the fields you have filled-in, so you can start over.
> Once Saved as PDF you will not be able to make changes.
> Upon clicking 'Save as PDF', a PDF copy of your form will be downloaded, which can be printed, emailed or uploaded to another application."

Some forms also have a link to a Mass.gov page with more information, but the pages I looked at have only general information rather than specific information for how to fill out a particular form.

The Trial Court has created an HTML5 form with "...tips for using Court online forms, and recommended best practices and standards for creating forms":
https://courtforms.jud.state.ma.us/publicforms/Court_Online_Forms_Guide.xdp

Appendix 5: Examples of Self-Help Materials

There is a sample of self-help material available from this "Courts Self-Help" page that is used frequently by litigants: https://www.mass.gov/topics/courts-self-help

Here are some specific examples:
- Small claims video guide: https://www.mass.gov/lists/small-claims-videos
- Responding to an eviction: https://www.mass.gov/service-details/respond-to-an-eviction-against-you
- Attending a virtual hearing: https://www.mass.gov/info-details/guide-to-virtual-hearings
- Accompanying instructions for forms. There are many examples of this on the Mass.gov, website. Here is the Affidavit of Indigency, as this is used across all Trial Court departments: https://www.mass.gov/doc/instructions-for-users-of-affidavit-of-indigency-and-its-supplement/download

Appendix 6: Participant Instructions

Thank you for coming in today and for agreeing to participate in this study.

My name is [Researcher's name]. During this session I'll be reading from a script to make sure I don't forget anything and that all participants hear the same instructions.

I'd like you to fill in some court forms so we can learn if any parts of them are confusing or hard to understand. Before we get started, however, I have a consent form that I'd like you to read, and then sign if you agree.

[alternative: I'd like you to read some information about legal procedures so we can learn...]

[Participant reads and signs consent form.]

As you are performing the tasks, please think out loud and tell me what is going through your mind. I want you to tell me about anything that is confusing, incomplete, unclear, missing, or incorrect. If you are quiet,   I may ask you what you're thinking about. Don't worry, as we'll practice this technique before we begin.

[When on paper or with the ability to mark up electronically]
In addition to thinking out loud and telling me what is going through your mind, I'd like you to underline, or circle, or add a question mark to any parts that are confusing or have some type of problem. If any part is especially clear, write a plus sign on the document.

[For forms we will probably not use the + to indicate a section that is clear. For self-help material we may use it.]

There is no such thing as an incorrect or wrong comment or answer. It's also important to realize that we're not testing or trying to evaluate you in any way. We are not testing you, we are testing the court forms and information. If you have trouble filling out a form, or understanding information about a legal procedure, then others will also likely have trouble or be confused, and that's why you're here. We need to find the parts that are confusing so we can improve them.

The entire study today will take about 45 minutes to an hour, and if you need a break at any time, just let me know. You can ask questions at any time, but I may want to delay answering some questions until the end of the session.

Before we begin, have you ever had to fill out [the form being tested; alternative: ...had to read any court information about ...]? Do you know anything about [whatever the legal procedure involves, for example, child support]?

Okay, let's get started. First, do you have any questions?

[The Researcher demonstrates the think-aloud technique and how to mark up a document with a non-court form, such as an Registry of Motor Vehicles form for Registration and Title Application]

Appendix 7: An Informed Consent Form

Here is a draft of an informed consent form that could be used for usability testing of court forms or documents.

Purpose of this Study

The Massachusett Trial Court is in the process of revising court forms to make them easier to understand and fill out. The court is also trying to improve information about legal procedures. Part of that process involves having people try to use the revised forms and information, and this study is part of that effort. We are asking you to fill out court forms and read information about court procedures so we can find problems with these materials and improve them.
.

Who is Conducting this Study?

The Trial Court has established a Forms Management Team to revise and evaluate forms and information. This team, led by xxx xxx, is working with a group of volunteers from Bentley University led by Dr. Demetrios Karis, a user-experience researcher. The researcher conducting the study may be either a court employee or a volunteer researcher.

Procedure

A researcher will give you instructions and examples, and you will then be asked to fill in a legal form or study some printed material. The researcher will ask you questions during and after you fill out the form. Then the researcher will give you a short questionnaire to fill out, pay you, and answer any questions you have about the study.

Recording

We would like to record the session, which will help us to take notes on your experiences while filling out the form. No one outside the research group will have access to these recordings and they will be deleted after the project is completed. If you do not want us to record, just say so at the beginning of the interview and we will take notes by hand during the interview rather than recording it.

Confidentiality

Your name will not be identified with the information we collect in any way. We will keep your name and the information you provide private and confidential. Nowhere in our report will we mention by name any of the people who participated.

Risks

There are no risks that we foresee in this study.

Breaks
If you need a break at any time, just let the researcher know.

Withdrawal
You can stop the interview and withdraw from the study at any time for any reason. If you withdraw, you will still be paid.

Payment
You will be paid xx in the form of a xxx gift certificate for your participation
or
There is no payment or compensation of any type for your participation.

Questions
You can ask questions at any time.

More Information
Contact [court person] and/or
Contact Dr. Demetrios Karis for more information: dkaris@bentley.edu

By signing this form, you are indicating that you understand and agree to the terms stated here and that you give the MA Trial Court permission to use your voice, verbal statements, and videotaped image for the purpose of evaluating and improving court forms and material. None of this material will ever be made public in any identifiable form.

Participant Signature: _____ Printed Name: _____ Date:_____
[In addition to DocuSign or some similar service, we could use Qualtrics to obtain a signature:
"The Signature question type presents survey participants with an entry box where they can draw their signature. On a desktop, they can use their mouse, and on a mobile device they can use their finger."]

Researcher Signature: _____ Printed Name: _____ Date:_____

Appendix 8: Interview Guide

What did you think of this [particular form or self-help material]? What's your general impression? Do you think you filled it out correctly and that it would be accepted by the court if you were to actually submit it?

Now I'd like you to please look back at the form [self-help material] and review with me why you marked up the form the way you did. [Alternative: and review some of the things you said while completing the form]

[Go through all the areas that the participant indicated were problematic in some way. In addition, the researcher will ask questions specific to the form tested that were prepared ahead of time.]

What else can you tell me about filling out this form [reading this material]?

[The researcher should ask additional questions based on what happened during the session or what the participant said.]

Appendix 9: Demographic Questionnaire

The court has already developed a few demographic questions for their Access and Fairness Survey, and we should use these questions in addition to developing new ones. It would be ideal to collect a complete set of demographic information, including age, sex and gender, ethnicity, marital status, location, education, employment status, income, and primary language. Here is a list of questions:
[include list of questions here]

Using a cognitive interviewing procedure, we will test and revise these questions.

Appendix 10: How to Test a Form: A Complete Example

1. The researcher greets the participant and reads the instructions (Appendix 6)
2. The researcher give the participant the informed consent form to read and sign (Appendix 7)
3. The researcher asks the participant about any previous experiences or knowledge with the form or legal procedure being investigated
4. The researcher demonstrates the think-aloud and mark-up technique using a non-court form such as a form for Registration and Title Application from the Registry of Motor Vehicles.
5. The researcher hands or shares (via screen sharing) a scenario with the participant and has them read it out loud. Along with the scenario is information the participant will need to fill out the form (e.g., names and addresses)
6. The researcher gives the participant a form to fill out (this form may have been sent to the participant ahead of time) or has the participant bring up an electronic version of the form and screen share
7. While the participant fills out the form, the researcher asks previously prepared questions at specific places in the form where there is the potential for confusion; alternatively, these questions can be asked after the participant has completed filling out the form
8. The researcher interviews the participant using the interview guide (Appendix 8)
9. The researcher has the participant fill out a short online demographic questionnaire (Appendix 9)
10. The research debriefs the participant (i.e., explains what the study is about), pays the participant, thanks the participant, and asks if there are any final comments or questions

A Scenario

For many issues of child custody, support, or visitation, a person will need to fill out a form titled, "Complaint for Custody-Support-Parenting Time Pursuant to L.L.c. 209C." Since this form is used for multiple purposes, we should probably develop different scenarios to test each of the main functions. Here is an example of a scenario for someone who wants to increase visitation time.

> "For the purpose of this study, please imagine that you are the father of John Jenkins, a six-year old boy who lives with his mother. You were never married to the mother, Martha Jenkins. You have been sending money to Martha to help support John, but she refuses to allow you to see him. You want the court to allow you to spend time with your son. Here is the form that you will need to fill out. Use your own name and the information provided here to fill out the form. While you fill out this form,

please read the instructions and choices out loud, and also tell me what you are thinking. If any word or question or part of the form is confusing, please underline or circle it. If any part is especially clear, write a plus in the margin."

The researcher will show the participant an example of a marked-up document and also have the participant practice this technique on part of another form. The researcher will ask the participant to use their own name, but the researcher will provide all other required information (e.g., the researcher would hand the participant a piece of paper with the name and address of the mother and the child). Alternatively, the researcher can tell the participant what his or her name is. We've used both variations and although it doesn't seem to make much difference, it may be preferable to have the participant fully pretend to be the person in the scenario and use a name we provide.

If the participant becomes stuck and asks for help, the researcher will typically respond, "What would you do if I weren't here and you were filling this out on your own?" If the participant says they would give up or contact the court for help, the researcher would note this but ask the participant to just go on and complete as much as possible.

After the participant has finished, the researcher would go through the form with the participant, asking about parts that were underlined or were associated with a plus in the margin. The researcher would also ask any pre-written questions about the forms that had not been covered.

Another Scenario
Here is another scenario for testing the Complaint for Custody-Support-Parenting Time.

"You have a five-year old son named Jonathan. You and Jonathan's mother never married, and the mother, Jennifer Alban, now has custody. You think she is doing a very bad job of taking care of Jonathan and so you want to have full custody.

Fill out the Complaint for Custody-Support-Parenting Time court form to get custody. You can find a PDF version of the form here:
https://www.mass.gov/doc/complaint-for-custody-support-parenting-time-cjd-109

Fill it out online and then print it out.

If you want, you can look at this page where there are some instructions:
https://www.mass.gov/service-details/instructions-complaint-for-custody-support-parenting-time-court-form

Here is information you will need to fill out the form.

Your name and address
Luke Crenshaw
45 Monroe Avenue
Barnstable, MA 02630

The name and address of the mother and child
Jennifer Alban
Jonathan Alban
224 Franklin Avenue, Apt 24C
Brewster, MA
02662

Jonathan was born on February 5, 2016. Although you are not listed on the birth certificate, you signed a statement saying that you are the father."

Appendix 11: How to Test Self-Help Material: An Example
See Appendix 10 above on How to Test a Form for the basic parts of an evaluation.

A Scenario
There is a Frequently Asked Questions (FAQ) document with 41 questions and answers about the Servicemembers Cases in the Land Court. The first FAQ is below. After the researcher welcomes the participant and gives instructions and the informed consent form, the researcher would ask the participant if he or she has any prior knowledge of Servicemembers Cases. The researcher will then read a scenario to the participant. For example:

> "For the purpose of this study, please imagine that you were in the army
> for eight years. Before joining the military you bought a house, and you
> paid the mortgage regularly while in the army. You left the military nine
> months ago to take a civilian job. Unfortunately, the company you worked
> for went bankrupt, you lost your job, and you were unable to make the
> mortgage payments on your house. The bank has just started foreclosure
> proceedings to take your house because you are far behind on your
> mortgage payments. A friend suggests you take a look at the
> Servicemembers Cases in the Land Court and sends you FAQs. Here
> they are. [The researcher hands the paper FAQs to the participant, or has
> the participant open the FAQs on a computer and share the screen over
> Zoom.] Please read these FAQs one at a time. If a word, sentence, or
> paragraph is confusing, underline or circle it. If a part is very clear, write a
> plus on or near it. After you read each FAQ I'll ask you to tell me what it
> means and to point out to me parts that were confusing or difficult to

understand. You can also give me feedback while you are reading the FAQ if you want."

The researcher will then show the participant an example of a marked-up document and also have the participant practice this technique on two paragraphs from another self-help document.

After the participant reads and marks up each of the FAQs (the first is below), the researcher will first ask the participant to explain, in their own words, the gist of the answer to the question. The researcher will next ask the participant to explain further why the marked parts were confusing or unclear. The researcher will also have a prepared set of questions about the FAQs in case the topics they cover don't arise naturally. For example, "junior lienholders" are mentioned in another Servicemembers FAQ, and if the participant does not underline this term the researcher might decide to ask the participant to explain who junior lienholders are[13].

1. What is a "Servicemembers" case?
"Under the federal law known as the Servicemembers Civil Relief Act (SCRA), a mortgage holder (also called a "mortgagee," usually a bank) cannot validly foreclose on a mortgage taken out before a servicemember's military service began, if the foreclosure is made during or within one year after the period of military service, unless a court issues an order allowing the foreclosure to proceed. In order to comply with this federal law, the bank (or other holder of a mortgage, trust deed, or similar security) files a "Servicemembers" case in state court against all owners of a property. The Servicemembers case is not a foreclosure of the property itself. In a Servicemembers case the only issue to be decided by the court is whether the owners of the mortgaged property are entitled to the benefits of the SCRA."

The first sentence of this example provides too much information, has too many conditional clauses (if…, unless…) and at over 60 words is too long. For most people it will probably require multiple rereadings. Ideally, sentences like this will be revised before testing starts. In some initial testing I did on this passage, the participant, who was highly educated and had some experience with legal forms and language, actually became agitated reading the first paragraph and used a curse word to express his frustration. I am happy to report that the FAQ on Mass.gov is somewhat improved:

---

[13] There is always judgment involved in moderating a usability session. Given how the participant responds to a particular FAQ, the researcher will decide which of the pre-written questions are appropriate to ask.

Appendix 12: A Categorization of Court Forms

Slightly different scenarios and testing procedures will be needed for different types of court forms. The essence of a good usability test is creating a situation for the participant that is as close to the natural situation as possible. For a form that needs to be submitted in the middle of a long and complex legal procedure this may not be easy to do. The litigant who will actually fill out this form will have previously filled out other forms for the same case and will have a variety of information and knowledge from the previous involvement with the case. To accurately evaluate this form we will thus need to provide the participant in our study with some of the knowledge and prior information that the litigant would have. This might include previous completed forms or information that was previously submitted to the court. It will thus be useful to know which forms are typically filled out at the beginning of a legal process and which forms usually come into play after a legal procedure has already started. In addition, some forms are typically submitted to the court along with other forms (e.g., financial statements, or the Affidavit of Indigency, which is filed with another complaint or petition). It will also be useful to know which forms come with instructions (typically printed on the back of the form) and which forms are associated with separate documents that include instructions.

My initial proposal is to categorize a form on three dimensions: whether instructions are included, whether the form can "stand alone" or is used after or in conjunction with other forms, and whether the form is short or long.

Instructions
1. No instructions on form and no instructions in a separate document
2. No instruction on form but there are instructions in a separate document
3. Instructions on the form and there is also a separate document with instructions
4. Instructions on the form but no separate document with instructions

Stand Alone
1. The form can stand alone. That is, a person could fill out this form without having previous knowledge about the case or having filled out previous forms
2. The form can not stand alone. That is, it is submitted along with another form, or was preceded by other forms that the litigant would have filled out and submitted to the court

Length
1. The form is short, defined as two pages or less
2. The form is long, defined as longer than two pages

Court Forms by Topic are presented at https://www.mass.gov/court-forms-by-topic. This page has links to 36 topics and most of the linked pages include multiple forms. I do not know what percentage of the total court forms are included on these pages.

Appendix 13: "Legalese" and Plain Language

Information and guidelines on plain language can now be found in multiple places, including the federal government, the National Center for State Courts, and the Self-Represented Litigation Network. See, for example, the following sites:

- https://www.plainlanguage.gov/
  - https://www.plainlanguage.gov/resources/content-types/legal-profession/
- https://www.ncsc.org/information-and-resources/trending-topics/trending-topics-landing-pg/what-is-plain-language-in-the-courts
- https://www.srln.org/node/150/srln-brief-plain-language-resources-100-access-srln-2015
- https://www.archives.gov/federal-register/write/legal-docs
  - https://www.archives.gov/federal-register/write/legal-docs/clear-writing.html

In addition, Clarity is directly relevant: "Clarity is the largest international plain language organization. It gathers professionals who are committed to promoting plain legal language and legal design." Their website has a lot of useful information: https://www.clarity-international.org/

Jury Instructions

In detailed and thoroughly documented experiments, Charrow and Charrow (1979) tested and confirmed three hypotheses: "(1) that the standard jury instructions used in this study -- when viewed as discourse -- are not well understood by jurors; (2) that certain linguistic constructions are largely responsible for the incomprehensibility; and (3) that if the problematic linguistic constructions are appropriately altered, comprehension will dramatically improve, notwithstanding the 'legal complexity' of any given instruction" (p. 1358). The jury instructions were written, although they were presented auditorily to the study participants. Given that, "Certain grammatical constructions and discourse structures found in the jury instructions appear to be recurring elements in legal language," I expect the findings to generalize to written material on forms, their instructions, and self-help material.

Charrow and Charrow measured the following linguistic constructions:
1. Nominalizations: This occurs when a noun has been constructed from a verb
2. Prepositional Phrases -- "as to.": these phrases can seem vague

3. Misplaced Phrases: Phrases inserted into the midst of otherwise normal clauses
4. "Whiz," and Complement Deletion: Subordinate clauses that are missing relative pronouns (that, which, who, etc.)
5. Lexical Items: The use of technical terms
6. Modals: The use of modal verbs such as "must," "should," and "may"
7. Negatives: The use of negatives in a sentence
8. Passives: Passive constructions or sentences
9. Word Lists: Situations in which several words are used when one will do
10. Discourse Structure: The organization of the information and individual sentences
11. Embeddings: The use of subordinate clauses within one sentence.

After modifying the instructions to reduce or improve the linguistic constructions above, Charrow and Charrow tested the modified instructions and found that the main points were much easier to understand, recall, and then repeat back while paraphrasing the instructions.

Other Research and Commentary

Redish (2010) provides an excellent history on evaluating the usability of documents, and also discusses the intertwined history of technical communicators and usability specialists. She also makes the important point that plain language needs to be about much more than short sentences and small words, but rather, "...plain language as UX [user experience] where people can (1) find what they need, (2) understand what they find, and (3) act appropriately on that understanding."

Dyer, Charles R.; Fairbanks, Joan E.; Greiner, M. Lynn; Barron, Kirsten; Skreen, Janet L.; Cerrillo-Ramirez, Josefina; Lee, Andrew; and Hinsee, Bill (2013) "Improving Access to Justice: Plain Language Family Law Court Forms in Washington State," Seattle Journal for Social Justice: Vol. 11 : Iss. 3 , Article 10. Available at: https://digitalcommons.law.seattleu.edu/sjsj/vol11/iss3/10

Mazur, B. (2000). Revisiting plain language. Technical Communication, 47, 205.

Passera S. Flowcharts, Swimlanes, and Timelines: Alternatives to Prose in Communicating Legal–Bureaucratic Instructions to Civil Servants. Journal of Business and Technical Communication. 2018;32(2):229-272. doi:10.1177/1050651917746459

Appendix 14: Readability Formulas

I advise against using readability formulas to calculate reading grade level. As Jarett & Redish (2019) write, "Readability formulas are easy to use and give you a number. This combination makes them seductive. But a number isn't useful if it isn't reliable, valid, or helpful." In fact, readability formulas are not reliable, they are not valid, and they are not helpful. The Massachusetts Trial Court Readability Guidelines are excellent, except for the first, which recommends using Word or another tool to determine reading level, and to aim for fifth grade literacy or lower. As Jarett & Reddish point out, "*Grade level* is a meaningless concept when writing for adults. What we really care about -- and what modern literacy assessments look for -- is *functional literacy*. Can adults understand what they are reading so they can do the tasks they need to do to find and keep jobs, take care of themselves and their families, and so on?" I would add, of course, can they understand the court's self-help material and successfully fill out legal forms without assistance.

Lewis (2006) found that common readability measures did not correlate significantly with the rated clarity of hardware and software documentation or measures of user satisfaction. The Cloudiness Count, however, did correlate significantly with these measures. The Cloudiness Count is based on the number of verbs in passive voice plus the number of "empty" words.

Charrow and Charrow (1979) conducted experiments in which they recruited people who had been called for jury duty and had them paraphrase jury instructions. They found no relationship between sentence length and comprehension. There was also no relationship between Flesch readability scores and the performance of participants, where performance was measured by the recall of the most essential units of the instructions. Charrow and Charrow rewrote and improved the jury instructions. They then found that, "the scores using a readability formula gave erroneous information regarding the actual comprehensibility of half of the instructions" (p. 1341). In some cases, where they confirmed that the rewritten instructions were superior based on improvements in comprehensibility, the Flesch scores were actually worse than the originals.

Crossley, Skalicky, & Dascalu (2019) review the literature on readability formulas and conclude that although they are widely used, they lack construct validity, and although they are predictive of the data sets on which they have been trained, they do not do well on new data sets. I'm sure it is possible to develop better automated measures of readability, but they will need to be much more complex than current formulas, and based on more than the number of syllables per word and sentence length. For example, Crossley et al. (2019) used a large corpus of crowdsourced judgments of

comprehension and reading speed (from Amazon's Mechanical Turk) and then developed "new readability formulas based on advanced NLP [natural language processing] tools for both text comprehension and reading speed. These formulas, based on linguistic features that better represent theoretical and behavioural accounts of the reading process, significantly outperformed classic readability formulas." Crossley et al. created two models, one for reading comprehension and one for reading speed. They found that, "Thirteen variables were significant predictors of the text comprehension ratings including variables related to lexical sophistication (word age of acquisition, frequency, imageability and character entropy), n-gram features (bigram range, trigram proposition scores and trigram TTR), cohesion (lexical overlap and the paragraph and sentence level) and sentiment (positive adjectives)" (p. 12).