

## **Collaborative Data Collection**

Jacqueline DaVania<sup>1</sup>, Kimberly A. Barchard<sup>1</sup>, and Wind Goodfriend<sup>2</sup>

1. University of Nevada, Las Vegas

2. Buena Vista University

### **Abstract**

Getting more people involved in data collection allows researchers to collect higher quality data from larger samples. Researchers can spread the time and cost of data collection among the entire research team. Because a team of researchers can afford more time and money than a single researcher can, collaborative data collection allows researchers to collect higher quality data. As well, collaboration allows larger samples. Sometimes our participants are difficult to recruit (such as people with bipolar disorder) and sometimes participants are free and easily accessible but in limited supply (as in the local Subject Pool). By collecting data in multiple locations, researchers can increase their sample sizes. This results in higher power, allows the use of multivariate statistical techniques, and allows the researcher to divide participants into subsamples.

There are three methods of collecting data at multiple sites. First, researchers can use traditional paper-based data collection. All materials are developed at one location, and distributed to the data collection sites. Second, researchers can use traditional computer-based data collection. Study materials are developed using programs such as Eprime and distributed to all data collection sites. In both of these methods, researchers send the data to a single location for analysis. Third, researchers can create a web site that contains all study materials, and participants from each institution access that same web site. The data is instantly sent to a single database for the entire project. While the first two methods of collaborative data collection are straight-forward, with testing materials developed using traditional techniques, putting a study on the Internet requires the researcher to learn how to create a web site that collects data.

There are two types of computer programs that can help researchers collect data through a web site. The first is a program designed specifically to create online forms. This includes SurveyMonkey, Event Handler, and SurveyWiz. The second is a general web development program, such as FrontPage or Dreamweaver. General web development programs are more flexible and allow researchers to program features such as skip questions, but require more knowledge and programming skill on the part of the researcher.

### **Introduction**

#### **Advantages of Collaborative Data Collection**

Getting more people involved in data collection allows researchers to collect data from larger samples. At each location, having more research assistants involved in the data collection, scoring, and data entry allows researchers to collect more data in a shorter period of time. However, usually each location has a limited supply of the required type of participant. Sometimes participants represent a rare or difficult-to-recruit population, such as infants, people with bipolar disorder, or skydivers. Sometimes participants are free and easily accessible but in limited supply, such as the local Subject Pool or high school classes. Therefore, if data can be collected in multiple locations, this can increase the sample size. Larger sample sizes result in higher power and more precise statistical estimates. Larger samples also allow the use of multivariate statistical techniques such as structural equation modeling. Finally, larger samples allow the researcher to divide the participants into sub-samples, to ensure that results hold true for different types of people. Alternatively, if a specific sample size is required, researchers can obtain that sample in a shorter period of time by collecting data at multiple sites.

Similarly, collaborative data collection allows researchers to collect higher quality data. Rather than using a short, quick test, the research team can use a longer, higher quality test, because the time required for data collection will be spread between many researchers. As well, rather than using an inexpensive test, the research team can use a more expensive test, and distribute the cost between all team members. Finally, rather than using measures that are easy and fast to score, they can use measures that are time consuming or expensive to score, such as second-by-second ratings of video-taped interviews.

Thus, a team of researchers can collect higher quality data from larger samples than an individual researcher can. The purpose of this presentation is to discuss different methods of collecting data collaboratively.

#### **Methods of Collaborative Data Collection**

There are three methods of collecting data from multiple locations: traditional paper-based data collection, traditional computer-based data collection, and online data collection.

##### **Traditional Paper-Based Materials**

First, data can be collected at multiple locations using traditional paper-based materials. The study materials can be developed in one location and then distributed to all data collection locations. If the materials include specific pieces of paper (such as item booklets or Scantron forms) then these pieces of paper can be sent to all sites. On the other hand, if the materials can be photocopied, the electronic files can be sent as email attachments, and then each location can make their own photocopies. After participants have completed the study, these pieces of paper can then be mailed to a single location for data entry, or each site can do their own data entry. If each site does their own data entry, it is probably best if the database is standardized for all locations. It is also possible to

scan in the responses and send these pictures of the data through electronic files.

One advantage of collecting data this way is that researchers can develop the study materials in the same way as they do for data collection they conduct on their own. They do not need to learn any new skills. However, collecting data this way is also challenging: researchers need to standardize the study procedures across the multiple locations. Researchers should develop a script for how the study administrators interact with participants. A script should be used any time there are multiple study administrators, but becomes more critical when the administrators have limited or no direct contact with each other. In addition, researchers may want to specify aspects of the physical setting, such as the size of the desk, the presence of windows, the presence of other people in the room, etc. When data are collected at one institution, these aspects can be standardized by collecting data in just one room (or just a handful of similar rooms). But when data is collected on paper at multiple locations, the physical rooms may vary tremendously, and so a discussion about standardized locations may be helpful. Finally, it is also possible to record variables that may affect the results, and to test for effects of these variables. For example, if one location resulted in different results than another, this problem could be solved in two ways. First, all analyses could statistically control for location. Second, the difference could be examined from a meaningful point of view – is there theoretical interest in how location might affect the outcomes? Are Midwesterners somehow different from Californians? Either way, these variables can be accounted for and can add a new dimension to the study.

### **Traditional Computer-Based Materials**

Second, data can be collected at multiple sites using traditional computer-based materials. Stimuli can be created using computer-based software such as Eprime or MediaLab. These stimuli can then be distributed to each of the data collection sites. After data collection is complete, the data files can be sent to one location for analysis. If each location has one or two computers with Eprime, then by collecting data at 5 or 10 institutions, data collection can be completed much faster.

In order to implement this type of data collection, though, each data collection site must have the relevant computer program. For example, in the example above, each data collection site must have EPrime and appropriate computer hardware to administer the study. In addition, the computer hardware and physical arrangement may need to be standardized. Often the monitor, keyboard, and computer speed will need to be the same at all locations, to ensure comparable data.

Both of the above two methods of collaborative data collection (traditional paper-based data collection and traditional computer-based data collection) are straight forward. The researcher develops the study materials using the same methods they would use if they were collecting data by themselves. To allow collaborative data collection, they only need to distribute these materials to the other data collection sites. The primary disadvantage of these methods is that they require standardization of the administration procedures. The researcher(s) who designs the study must ensure that participants are given the same instructions and the same time-limits at all data collection sites. As well, they must ensure that the materials are given in the same order at all sites and that all other procedures are also standardized. Because of this, the computer-based methods have an advantage over the paper-based methods, because the principle investigator can ensure that time limits, instructions, and the order of the stimuli are identical. The more procedures are implemented by the study administrator, the harder it is to control and standardize the administration.

### **Internet-Based Data Collection**

The third method of collecting data at multiple sites is to put all study materials on the Internet, and have the participants at all sites accessing that one site. In this way, the participants are given identical consent procedures, identical stimuli, identical time limits, and identical debriefing. Thus, all procedures can be standardized. This is particularly helpful for collaborative research studies, because otherwise it can be difficult to standardize administration procedures across multiple institutions. Furthermore, the data can be instantly assembled into a single database. This is particularly useful for collaborative data collection, because no additional step is needed to assemble the data into one location, and furthermore, the researcher is assured that data are entered in identical ways at all sites.

In some cases, the participants will be able to access the Internet website by themselves. This is the fastest method of collecting data, because appointments are not needed. Participants can access the study at any hour of the day or night, and on any day of the week, because they are using their own computers. This allows the researcher to collect a lot of data in a short period of time. One disadvantage of having participants access the web site by themselves is that some degree of standardization is lost. The participant might access the study when they are in a quiet location or a noisy one. They might be alone or talking to others. They might have a large monitor or a small monitor. A second disadvantage of having participants access the web site by themselves is that this makes it impossible to maintain test security. Many test publishers will refuse researchers permission to use a published test in an online study, if participants will have unsupervised access to the test materials. See Barchard and Williams (in press) for a discussion of copyright issues in online studies.

In other cases, the participants will meet with the researcher in individual or group appointments. The participant will be asked to sit at a computer where the researcher has loaded the study web site. The participant then completes the study on their own. The debriefing can be done by the web site or by talking with the study administrator. This type of arrangement may be necessary in some circumstances. For example, this might be necessary to ensure that participants are debriefed, which is essential if the study involves deception. Or appointments might be necessary to ensure standardization of the environment and the computer equipment. For

example, studies that measure reaction time should standardize the speed of the computer. Or appointments may be necessary if the study is using copyrighted tests, because this allows the researchers to put the test materials on a password-protected site and keep that password secret.

If individual appointments are needed, it is still beneficial to have most of the study materials on web sites for three reasons. First, this facilitates standardized procedures across the institutions that are participating in the study. Second, this once again ensures that all data are entered into the same database, using identical entering procedures. Third, the benefits of computer-administration can be obtained without each data collection site needing to have specialized software. The web site can be created by one researcher, and the only software that is needed to run them is an Internet browser.

## Two Methods of Implementing Internet-Based Data Collection

In the rest of this presentation, we will discuss two methods of creating a web site that collects study data: programs that create online surveys and general web development programs.

### *Programs that Create Online Surveys*

There are many programs that were designed specifically to create online surveys and web experiments. See Table 1. These programs are generally pretty easy to use. The researcher tells the program what items to include, and whether to use radio buttons or check boxes.

Table 1  
Programs Specifically Designed to Create Online Surveys and/or Web Experiments

Program	Developer	Available from
EventHandler 4.0a	UbiDog Productions	<a href="http://www.event-handler.com/">http://www.event-handler.com/</a>
LiveCycle Designer	Adobe	<a href="http://www.adobe.com/products/server/adobedesigner/">http://www.adobe.com/products/server/adobedesigner/</a>
mrInterview	SPSS	<a href="http://www.spss.com/mrinterview/">http://www.spss.com/mrinterview/</a>
ERM Feedback	Vovici	<a href="http://www.vovici.com/">http://www.vovici.com/</a>
SurveyMonkey	SurveyMonkey.com	<a href="http://www.surveymonkey.com">http://www.surveymonkey.com</a>
The Survey System	Creative Research Systems	<a href="http://www.surveysystem.com/">http://www.surveysystem.com/</a>
SurveyWiz (free)	Birnbaum	<a href="http://psych.fullerton.edu/mbirnbaum/programs/surveyWiz.htm">http://psych.fullerton.edu/mbirnbaum/programs/surveyWiz.htm</a>
WEXTOR (free for educational and non-commercial uses)	Reips & Neuhaus	<a href="http://psych-wextor.unizh.ch">http://psych-wextor.unizh.ch</a>
PsychData	PsychData, LLC	<a href="http://www.psychdata.com/">http://www.psychdata.com/</a>

To give readers a sense of what these programs are like, we will present an example of one of these programs. SurveyMonkey is a program that was designed specifically to create online forms. It is simple to use and gives point and click options. First the researcher creates the overall survey and selects options. There are options such as page numbering and item numbering, progress bars, page titles and navigation buttons. Figure 1 shows the survey design options in SurveyMonkey

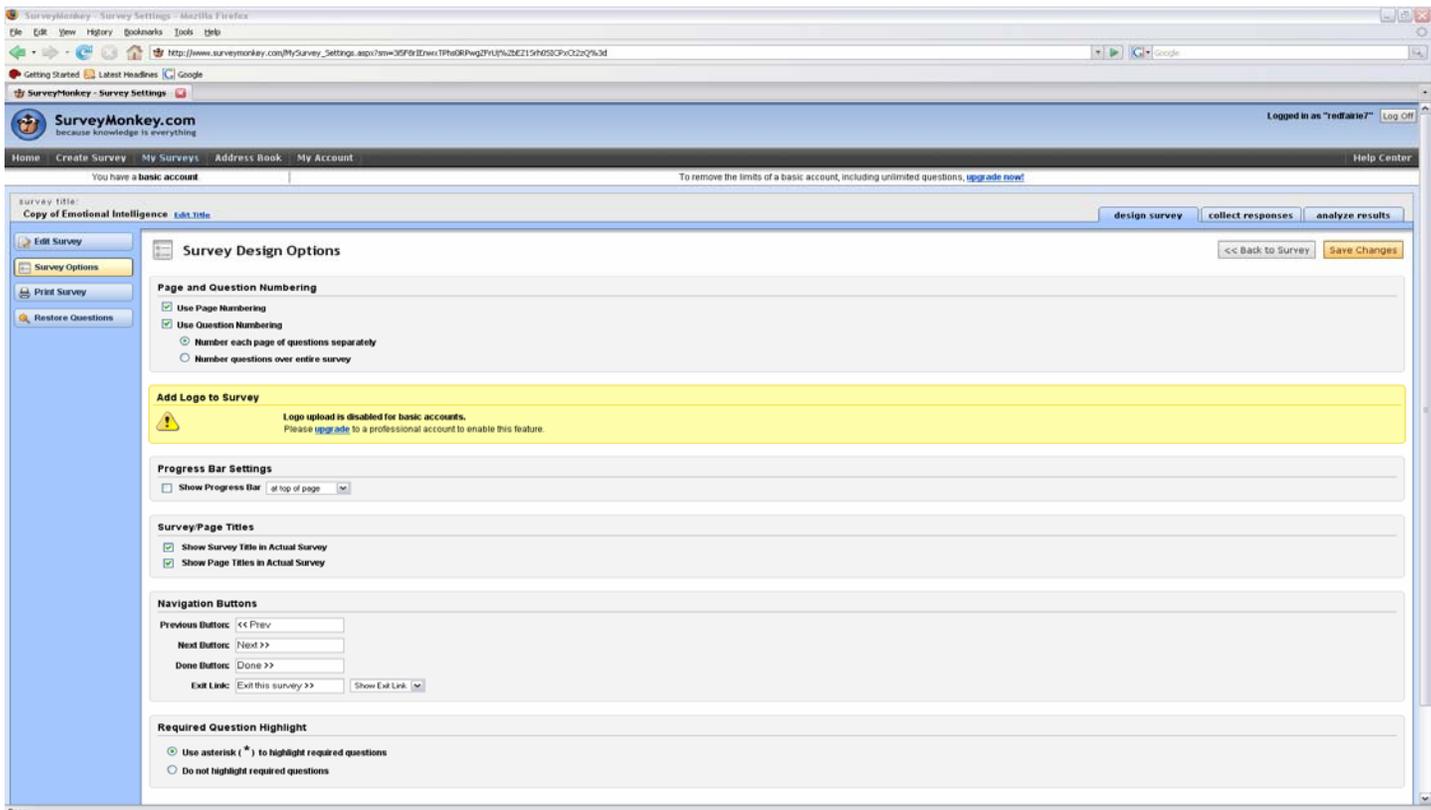


Figure 1  
Survey Design Options in SurveyMonkey

To add questions to the survey, click the Add Questions button. For each item, choose the type of question. SurveyMonkey has several item types, such as textbox, multiple choice, matrix of choices, images, and rating scales. Finally, enter the question and response options. See Figure 2.

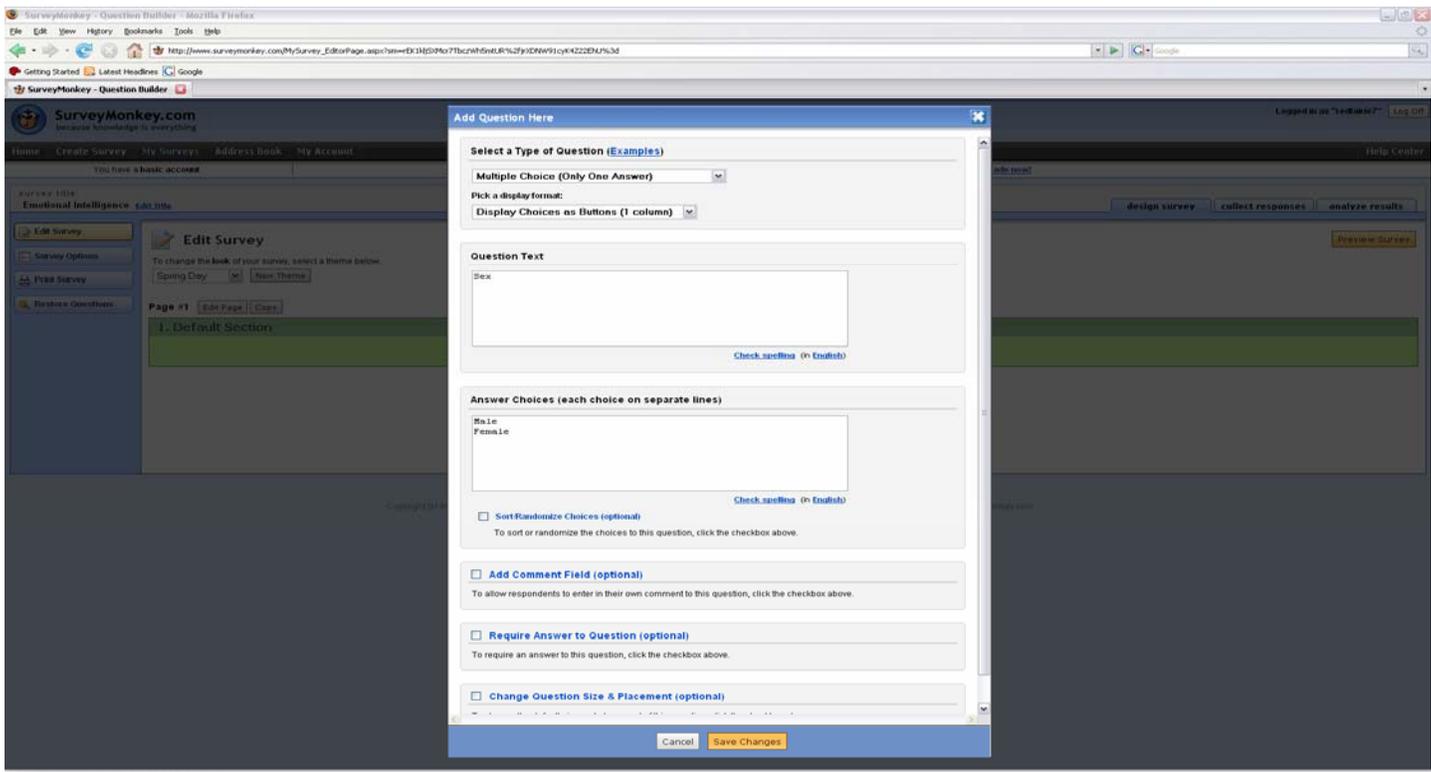


Figure 2  
Adding a Question in SurveyMonkey

When all of your questions are entered, you can administer the survey in three ways. One option is to print your survey and distribute it on paper. The second option is to create a set of emails that advertise your survey, and provide a direct link to the online form. The third option is to add a link to your survey onto an existing website. Thus, SurveyMonkey provides researchers with two different ways of advertising and administering these surveys online.

In addition, SurveyMonkey provides tools for analyzing responses after data collection is complete. Thus, this program may be particularly attractive to researchers (such as students) who do not have ready access to statistical packages such as SPSS.

### General Web Development Programs

General web development programs are more flexible than programs that were designed just to create online forms. They can be used to create online forms, as well as regular web pages that don't have any online forms. Many of these programs are easy to use and some are free. See Table 2 for a listing of some of these programs.

Table 2

#### Easy-to-Use Web Development Software

Program	Developer	Available from
Amaya (free)	W3C	http://www.w3.org/Amaya/
Authorware 7	Adobe	http://www.adobe.com/products/authorware/
Dreamweaver CS3	Adobe	http://www.adobe.com/products/dreamweaver/
1st Page 2006 (free)	EvrSoft	http://www.evrsoft.com/
Flash 8	Adobe	http://www.adobe.com/products/flash/
Expressions	Microsoft	http://www.microsoft.com/expression/
FrontPage 2003	Microsoft	http://www.microsoft.com/frontpage/

In the rest of this presentation, we will focus on one general web development program: Dreamweaver. Dreamweaver basically looks like a word processing program. See Figure 3. The user can see the actual words and figures that will be shown on the web pages. The technical term for this type of web development program is WYSIWYG. This is an abbreviation of the phrase "What You See Is What You Get".

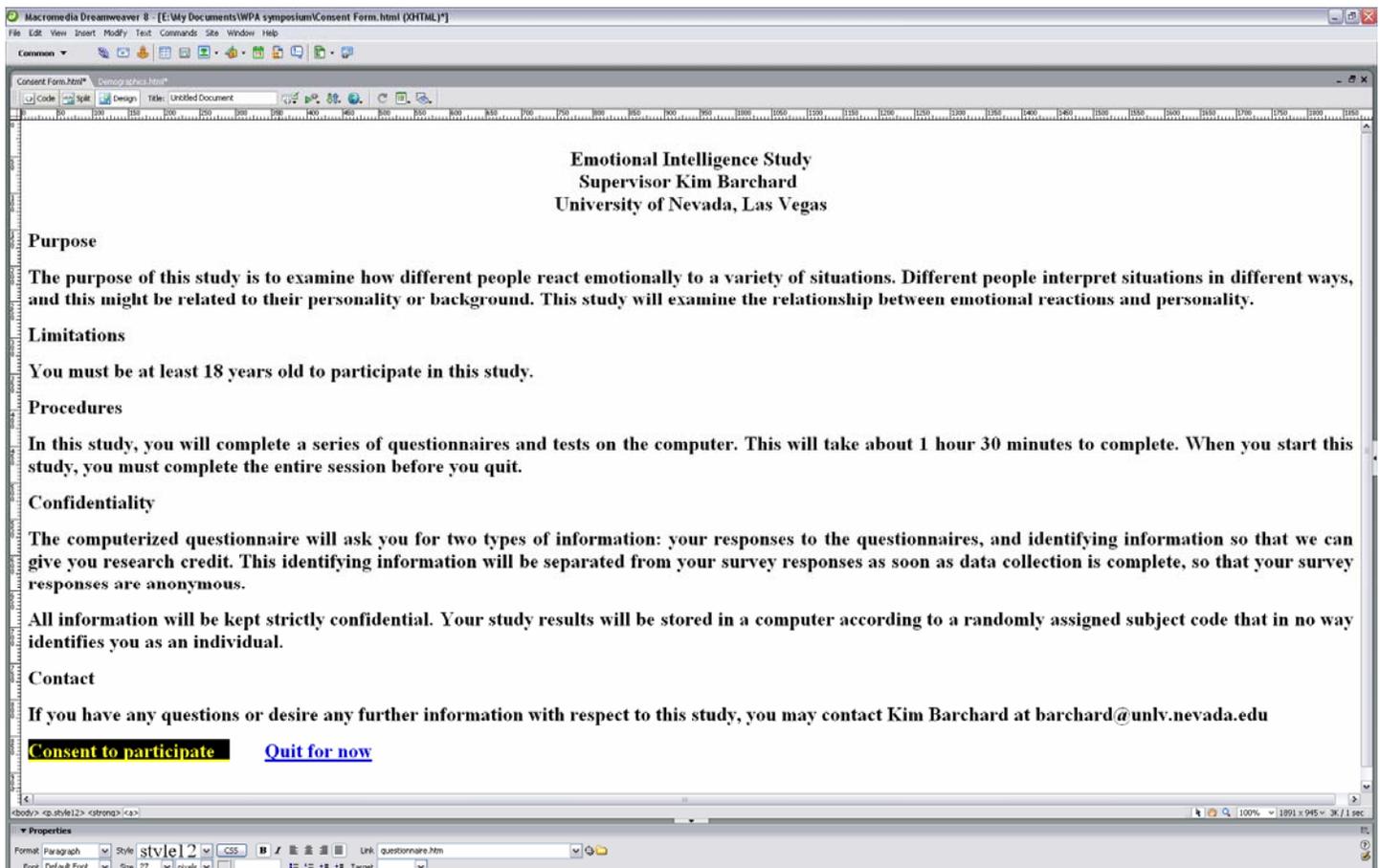


Figure 3: Demographics  
Creating a Web Page in Dreamweaver

Creating web pages in Dreamweaver is easy. To create a new web page, just click on the File menu, click New, and click Create. Type the content you want, and format it the way you want. There are buttons at the bottom of the screen on the Properties Toolbar to make the text bold, italics, and centered, and to change the font style and size. To save the web page, click File, Save.

When you have created several pages, such as a consent form page, questionnaire page, and debriefing page, you will want to link them together. First, select the words you want to use as the link. Then type the name of the next web page into the Link box at the bottom of the screen. See the bottom of Figure 3.

Creating a questionnaire in Dreamweaver is also easy. First create a basic web page or open a new blank page. You can put the name of the questionnaire at the top, if you'd like. Next, insert a blank form, which will contain the questionnaire items. To insert a form, click Insert, Form and click the type of form you want. This will show up as a small red box that you can type things in. You can make it bigger by adding a few blank lines. See Figure 4.

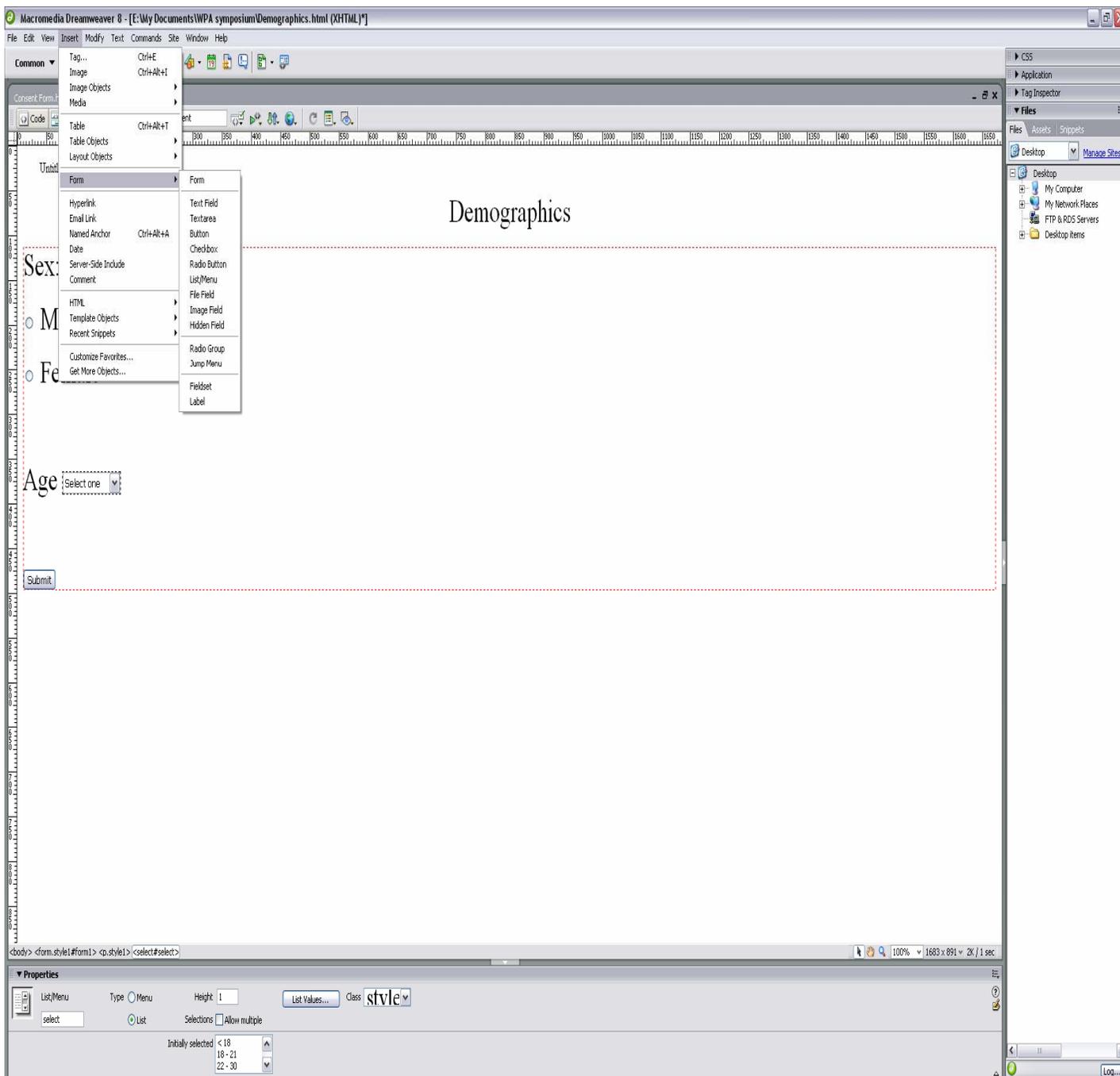


Figure 4  
Creating a Blank Form in Dreamweaver

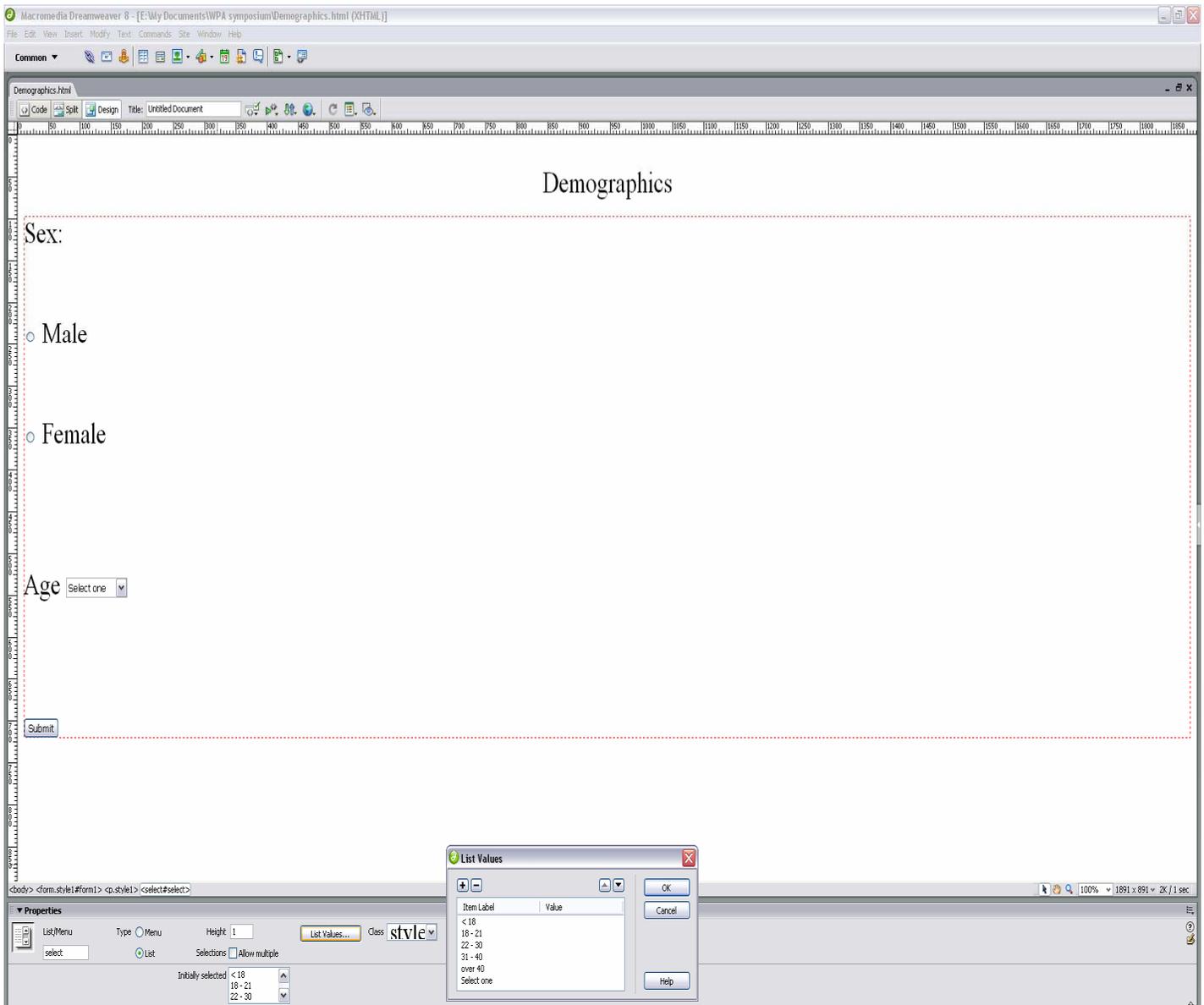


Figure 5  
Creating a Radio Button in Dreamweaver

Radio buttons are very useful. With a set of radio buttons, all the options for a question are shown at once. For example, if you want to ask someone what Sex they are, you can have the options of Male or Female. To create the Radio buttons, click Insert, Form, Radio Button. Type the label you want displayed next to the radio button: in this case, Male or Female.

Drop-down menus are also very useful, especially if you have many possible responses. These works were for variables such as Age and Ethnicity. To create a drop-down menu, click Insert, Form, List/Menu. Type the label you want displayed next to the drop-down menu, in this case, the word Age. Next, to add the response options, click List Values. Click the + symbol to add response options. In this case, we had several different age ranges. In Figure 5, the List Values window is on the bottom middle.

Adding pictures is also easy. Just click Insert, Image, and then select the file where your image is. This can be useful if your stimuli consist of pictures or videos.

After you have created all of your items, you need to insert a Submit button, so that participants can send the data to you. Click Insert, Form, Button, OK.

The last step is to tell Dreamweaver where to send the data. The easiest method is to have the data emailed to you. In the Form Properties window, in the Action box, type "mailto:" and follow this with your email address. See Figure 5. A more elegant method is to send the data to a form processor. See [http://php.resourceindex.com/Complete\\_Scripts/Form\\_Processing/](http://php.resourceindex.com/Complete_Scripts/Form_Processing/) for examples of form processors.

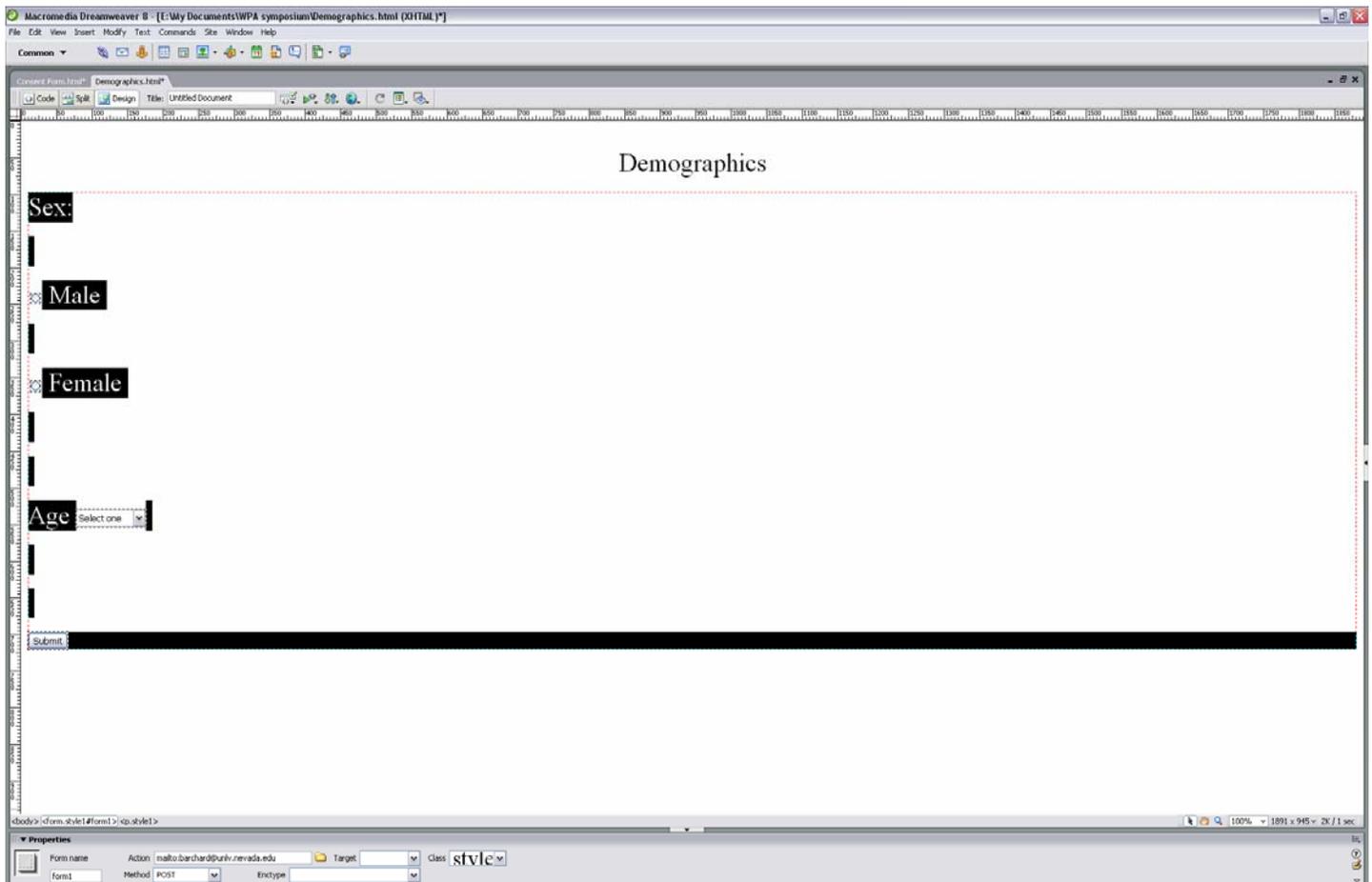


Figure 6  
Sending Data to Email

Dreamweaver can be purchased through the Adobe website. See Table 2. And you can try out the latest version of Dreamweaver for 30 days for free. Go to <http://www.adobe.com/products/dreamweaver/> and click on "Download Free Trial".

### Conclusions

There are many ways that researchers can collaborate on data collection. First, they can just design the study together, and then collect data at one site. Second, they can collect data at multiple sites, using traditional paper or computer-based data collection methods. Third, this presentation has highlighted methods of collecting data online, either using programs that were designed specifically to create online surveys or general web development programs such as Dreamweaver.

Whatever method is used, there are several advantages of collaborative data collection. First, researchers can collect higher quality data, because the time and expense of data collection is spread between many people. Second, researchers can collect larger samples more quickly. This increases reliability, effect size, power, and significance. Third, this allows researchers to collect more diverse samples, which can then be analyzed by subgroups to ensure that results and conclusions are applicable to each type of participant. All of these allow researchers to collect data that is high quality, important and meaningful, while still being fast, easy, and cheap. By collaborating, data quality does not have to be sacrificed to practical issues.

### References

Barchard, K.A. & Williams, J. (in press). Practical advice for conducting ethical online experiments and surveys for United States psychologists. *Behavior Research Methods*.

### Acknowledgements

We would like to thank the following people, who provided feedback on earlier drafts of this presentation:  
Stephanie Rojas, University of Nevada, Las Vegas