

9 August, 2007

Status
Final

Revision
V1.0

Document File Name
*Bluetooth Secure Simple
Pairing Usability Metric
Whitepaper*

Document Owner
Usability Expert Group

E-mail Address
usability-
main@bluetooth.org



BLUETOOTH SECURE SIMPLE PAIRING USABILITY METRIC WHITEPAPER

Test What You Build

ABSTRACT: This white paper provides a set of Usability Metrics and the instructions on how to apply them to the *Bluetooth Secure Simple Pairing* process. While this document is specifically targeted at *Bluetooth Secure Simple Pairing* and using the metrics for this process, they are equally applicable to other usage scenarios.



Revision History		
Revision	Date	Description
R01	Sept 21, 2006	Initial Draft
R02	Sept 22, 2006	UEG face-to-face edits
R03	Sept 22, 2006	UEG face-to-face edits
R05	Oct 2, 2006	UEG conference call edits
D05_r06	November 8, 2006	BARB review edits.
D05_r07	November 29, 2006	Marketing review edits. Incorporated updated lexicon definitions and changes.
D05_r08	January 17, 2007	UEG face-to-face edits.
D09_r00	January 18, 2007	Final edits and UEG walkthrough.
D09_r07	March 21, 2007	Final edits and BARB / UEG walkthrough.
D09_r08	July 16, 2007	Editorial review
V1.0	August 9, 2007	Release & change to Bluetooth Secure Simple Pairing Usability Metric Whitepaper

Contributors	
Name	Company
Harry Zhang	Motorola
Kimberly Saucedo	Plantronics
Ryan Seick	Motorola
Niall Gillespie	CSR
Frank Armstrong	Kodak
Paul Wootten	TMTI
Jun Shen	AT&T
Yeesin Chan	Sprint
Emma Oivio	Nokia
Peter Hauser	Bluetooth SIG
Rob Franzo	Broadcom

DISCLAIMER AND COPYRIGHT NOTICE

THIS DOCUMENT IS PROVIDED "AS IS" WITH NO WARRANTIES WHATSOEVER, INCLUDING ANY WARRANTY OF MERCHANTABILITY, NONINFRINGEMENT, FITNESS FOR ANY PARTICULAR PURPOSE, OR ANY WARRANTY OTHERWISE ARISING OUT OF ANY PROPOSAL, SPECIFICATION OR SAMPLE. Any liability, including liability for infringement of any proprietary rights, relating to use of information in this document is disclaimed. No license, express or implied, by estoppel or otherwise, to any intellectual property rights are granted herein.

This document is for comment only and is subject to change without notice.

Copyright © 2001–2007 Bluetooth® SIG, Inc. All copyrights in the *Bluetooth* Specifications themselves are owned by Ericsson AB, Lenovo, Intel Corporation, Microsoft Corporation, Motorola, Inc., Nokia Corporation, and Toshiba Corporation. *Other third-party brands and names are the property of their respective owners.



CONTENTS

1	INTRODUCTION.....	5
2	BIBLIOGRAPHY.....	6
3	METRIC OVERVIEW.....	7
4	DEFINITION OF METRICS	8
	4.1 OVERALL TASK TIME	8
	4.1.1 SETUP TIME	8
	4.1.2 <i>BLUETOOTH</i> SECURE SIMPLE PAIRING TASK TIME	8
	4.2 ERROR COUNT.....	8
	4.3 FATAL ERROR(S).....	9
	4.4 TASK INCOMPLETION RATE	9
	4.5 INTUITIVENESS	9
	4.6 USER SATISFACTION	9
5	APPLICATION OF METRICS	11
	5.1 JUST WORKS TASK FLOW	11
	5.1.1 <i>BLUETOOTH</i> SECURE SIMPLE PAIRING TASK TIME	11
	5.1.2 ERROR COUNT EXAMPLES.....	11
	5.1.3 FATAL ERROR(S) EXAMPLES	12
	5.1.4 TASK INCOMPLETION RATE EXAMPLES	12
	5.1.5 INTUITIVENESS EXAMPLES	12
	5.1.6 USER SATISFACTION EXAMPLES	12
	5.2 CONNECT TASK FLOW	13
	5.2.1 OVERALL TASK TIME	13
	5.2.2 ERROR COUNT EXAMPLES.....	13
	5.2.3 FATAL ERROR(S) EXAMPLES	14
	5.2.4 TASK INCOMPLETION RATE EXAMPLES	14
	5.2.5 INTUITIVENESS EXAMPLE	14
	5.2.6 USER SATISFACTION EXAMPLES	14
	5.3 SEND / RECEIVE TASK FLOW	15
	5.3.1 OVERALL TASK TIME	15
	5.3.2 ERROR COUNT EXAMPLES.....	15
	5.3.3 FATAL ERROR(S) EXAMPLES	16
	5.3.4 TASK INCOMPLETION RATE EXAMPLES	16
	5.3.5 INTUITIVENESS EXAMPLES	16
	5.3.6 USER SATISFACTION EXAMPLES	16
	APPENDIX	17



List of Figures

FIGURE 1: JUST WORKS TASK FLOW	11
FIGURE 2: CONNECT TASK FLOW.....	13
FIGURE 3: SEND/RECEIVE TASK FLOW	15



1 Introduction

In June of 2006 the *Bluetooth* SIG formed the Usability Experts Group¹ with the goal:

“The Usability Expert Group’s goal is to bring usability experts, marketing, development, and customer service experts together in a collaborative forum to enhance Bluetooth usability across all Bluetooth wireless technology enabled products.”

This white paper describes the work done in the area of Usability Metrics by the group. To develop the metrics the team reviewed previous *Bluetooth* usability analysis² and data collected during the support of commercially deployed products^{3 4}. These data focused the group upon the pairing process as the area where most gain can be achieved in usability.

The focus of this white paper is on Usability Metrics. The Example *Bluetooth* Secure Simple Pairing Survey in the Appendix provides a set of quantifiable close-ended questions. There is value in using qualitative open-ended questions in usability surveys to gather emotional and abstract feedback from users on their experience yet such questions are outside the scope of this document.

During the work on metrics several design considerations were also identified. This white paper does not address design consideration as these are included in the *Bluetooth* Secure Simple Pairing Flow Diagram whitepaper.

While this document is specifically targeted at *Bluetooth* Secure Simple Pairing the metrics for this process are equally applicable to other usage scenarios.

NOTE: If there is a discrepancy between the definitions and procedures within this document and the underlying *Bluetooth* specifications, the *Bluetooth* specifications override any definition or procedure herein.

1 Bluetooth Usability Expert Group Overview Document

2 First-time Usability Testing for Bluetooth-enabled Devices

3 Wootten, Paul, TMTI, Presentation on their experience with usability issues in the marketplace, June 28, 2006.

4 Wootten, Paul, TMTI, Presentation on basic statistics of usability issues in the marketplace, July 27, 2006.



2 Bibliography

- [1] The following sources are cited as references as noted throughout the document:
- [2] *Bluetooth* Secure Simple Pairing Lexicon,
https://programs.Bluetooth.org/docman/DirInfo.aspx?gld=498&dir_id=2070
- [3] *Bluetooth* Secure Simple Pairing Whitepaper,
https://www.Bluetooth.org/foundry/adopters/document/Simple_Pairing_WP_V10
- [4] *Bluetooth* Usability Expert Group Overview Document,
https://programs.Bluetooth.org/docman/DocInfo.aspx?doc_id=41578&vld=43990&gld=90
- [5] First-time Usability Testing for *Bluetooth*-enabled Devices,
https://programs.Bluetooth.org/docman/handlers/DownloadDoc.ashx?doc_id=41199&vld=43257
- [6] First Connect FRD: First_Connect_FRD_V10r00.doc:
https://programs.Bluetooth.org/docman/DocInfo.aspx?doc_id=40874&vld=42721&gld=59
- [7] Howell, D. C., *Statistical methods for psychology*. Second edition. Boston: Duxbury Press. 1987.
- [8] Nielson, J., *Usability engineering*. Cambridge, MA: Academic Press Professional. 1993
- [9] Wickens, C. D., Gordon, S. E., & Liu, Y., *An introduction to human factors engineering*, 1998.
- [10] Wootten, Paul, TMTI, Presentation on their experience with usability issues in the marketplace, June 28, 2006.
- [11] Wootten, Paul, TMTI, Presentation on basic statistics of usability issues in the marketplace, July 27, 2006.



3 Metric Overview

System/device usability refers to the ease of use and intuitiveness of device user interfaces. User interfaces with a high level of usability are simple, user-friendly, and compatible with user mental models. According to Nielsen⁵, these devices will have the following characteristics.

- Learnability and efficiency: Users can complete goal-oriented tasks rapidly at the first use and subsequent uses.
- Memorability: Users do not need to re-learn the system operation after a period of non-use.
- Low Errors: Users should make few errors and can correct errors when they occur.
- Satisfaction: Users should be subjectively satisfied with using the system.

Errors are committed when user actions deviate from normal procedure and an incorrect decision is made. Two types of errors exist: fatal and non-fatal. A fatal error leads to immediate task incompleteness. A non-fatal does not immediately terminate a task.

Usability may be measured with a multitude of well-defined metrics⁶. These metrics will be defined in the next section. Usability has been measured using the following methods.

- Heuristic evaluation: One or multiple usability experts will analyze the characteristics of devices or systems to determine whether usability standards and guidelines have been followed.
- Usability testing/experiment: Disinterested participants randomly selected from a population are asked to pair two devices. Objective measures such as time and error will be captured.
- Focus groups: A group discussion around a predefined topic is initiated to assess first impressions and user reactions to particular products.
- User interviews: An interview session is conducted to acquire users' opinions about products, and their preferences, areas of difficulty, patterns of use, and areas of improvement.
- Surveys: A set of written, device-specific questions or scales is used to assess users' opinions about products, areas of difficulty, and areas of improvement.
- Psychometric questionnaires: Users are given a standardized set of questions and their responses are compared against norms derived from other users to similar products. For example, Software Usability Measurement Inventory (SUMI) has been used for software testing (<http://sumi.ucc.ie/>).

These methods are often combined to afford a better understanding of device usability. For example, a usability experiment may be performed on pairing a cell phone with a headset to acquire pairing time and errors, and after the experiment, a survey may be given to acquire users' opinions.

Not all methods are required for assessment of *Bluetooth* Secure Simple Pairing. Some methods may not be directly applicable to pairing. Because it is desirable to have a 5-point scale to assess pairing tasks, 3-point scales used in SUMI may not be suitable.

Products should complete functionality testing and interoperability testing prior to any usability testing. A baseline of product usability should be developed using the appropriate set of metrics to guide the improvement of product usability.

The experiments and surveys should use a sample that represents the user population of interest. Because there might be large individual differences among users, the sample size should be sufficiently large to afford adequate statistical power⁷.

5 Nielsen, J. (1993). Usability engineering. Cambridge, MA: Academic Press Professional.

6 Wickens, C. D., Gordon, S. E., & Liu, Y. (1998). An introduction to human factors engineering. Addison Wesley Longman, Inc.

7 Howell, D. C. (1987). Statistical methods for psychology. Second edition. Boston: Duxbury Press.



4 Definition of Metrics

4.1 OVERALL TASK TIME

Overall Task Time is defined as the total time spent to complete a given task. For a pairing task, it includes both the setup time and *Bluetooth* Secure Simple Pairing task time.

4.1.1 SETUP TIME

The setup time for a pairing task is defined as the time that a user takes to put devices into the searching or discoverable mode. The specific procedure for putting devices into the searching or discoverable mode will vary for different devices and user interfaces. Some devices may have the *Bluetooth* link on the main menu, and other devices may place the *Bluetooth* link several layers deep in the menu structure. The setup time will include the time that the user navigates through menu items such as settings, connections, *Bluetooth* links, and setup to put the devices into the searching or discoverable mode.

4.1.2 BLUETOOTH SECURE SIMPLE PAIRING TASK TIME

For the *Bluetooth* Secure Simple Pairing process, the start of the task time is when a user action initiates the pairing process that puts devices into the discoverability and searching mode. If a device enters the discoverability mode upon powering on, then this would be the starting point. The end of the task time is when the user receives confirmation on both devices that the task has been completed.

Note: The process of making a *Bluetooth* device discoverable should be simple and completed as quickly as possible.

Overall task time for the *Bluetooth* Secure Simple Pairing process should be shorter than the length of discoverability. While the goal is to make this as short as possible, for *Bluetooth* Secure Simple Pairing of two devices the maximum for overall task time should be 30 seconds⁸; whereas, discoverability should be at least three times the overall task time (greater than 90 seconds for *Bluetooth* Secure Simple Pairing devices including powering up).⁹

Task time is directly influenced by the number of steps, number of key/button presses, number of decisions and number of errors.

Overall Task Time should also be established for Legacy Pairing devices, *Bluetooth* Secure Simple Pairing devices and combinations of these pairing types.

An example of Overall Task Time:

- User selects “Send file” (on camera menu screen) to printer. This action begins the start task time. When the user confirms the codes on both devices and selects “OK”, then the process is ended.
- See Sections 5.1.1, 5.2.1, and 5.3.1 for examples of Overall Task Time metric application.

4.2 ERROR COUNT

Error Count is defined as the number of errors made by users. Errors are committed when user actions deviate from normal procedure and an incorrect decision is made.

Two types of errors exist, fatal and non-fatal. Fatal error examples would be typing incorrect passkeys or not being able to put the device in discoverable mode. Non-fatal error examples would be pushing a wrong button that has no major impact, accepting passkeys without reviewing/comparing, making wrong decisions or skipping a major step. If a user selects an option that is not within the steps of the original procedure and multiple steps are taken after this, only one error should be recorded for this.

8 First Connect FRD: First_Connect_FRD_V10r00.doc

9 *Bluetooth* Secure Simple Pairing Whitepaper



There are two ways to detect an error. The first is a manual or procedural error, which could be an observer-detected error. Examples of this are choosing the wrong thing or a copy/paste vs. cut/paste. The second way to detect an error is with software. These software-detected errors could be entering a wrong number or entering five digits instead of six.

To track procedural error count, developers can create tools to assist in evaluation.

See Sections 5.1.2, 5.2.2, and 5.3.2 for examples of Error Count metric application.

4.3 FATAL ERROR(S)

Fatal Error(s) are defined as the number of errors that can lead to immediate task incompleteness. Fatal Errors can cause user frustration, user dissatisfaction or even the return of devices to where they were purchased. Compilation of Fatal Errors across a representative sample of users and/or user attempts will typically be required for useful measure of this metric.

Examples of fatal errors would be typing wrong passkeys or not being able to put the device into discoverable mode. In the camera example, if the user accepts the code on the printer, but it is not the same code that was on the camera, the pairing is not successful¹⁰, which results in a fatal error.

See Sections 5.1.3, 5.2.3, and 5.3.3 for Fatal Error metric applications.

4.4 TASK INCOMPLETION RATE

Task Incompletion Rate is defined as percentage of users who do not complete the *Bluetooth* Secure Simple Pairing task within the discoverability window, and of users who choose to give up due to frustration.

The task may not be completed for many reasons. The following are examples of potential causes for users not completing the task:

- Users do not know the full procedure.
- Devices are not powered up.
- Devices are not in discoverable mode.
- Task completion time exceeds the length of discoverability.
- Devices are incompatible: *Bluetooth* Wireless Technology enabled headset & non-*Bluetooth* Wireless Technology enabled phone.
- Users pressed the wrong button.
- Users cannot determine what to select, etc.
- Out Of Band mechanism mismatch.

See Sections 5.1.4, 5.2.4, and 5.3.4 for Task Incompletion Rate metric applications.

4.5 INTUITIVENESS

Intuitiveness is defined as the ease of understanding the *Bluetooth* Secure Simple Pairing process with minimal confusion and without having to consult a manual. The target is for the user to rate highly on ease of completing the pairing process without having referred to the manual.

The more intuitive *Bluetooth* Secure Simple Pairing process results in lower overall task time, lower error count, higher task completion and higher user satisfaction.

See the Appendix for examples of user survey questions related to measuring Intuitiveness.

4.6 USER SATISFACTION

User Satisfaction is defined as what users think and how they feel about a product and its perceived quality. It is a composite measure of the above metrics as well as other measures. User Satisfaction for *Bluetooth* Secure Simple Pairing can be influenced by reliability, robustness, intuitiveness, ease of pairing instructions,

¹⁰ If there was a Man-In-The-Middle attack, the pairing would be viewed as successful from the devices and not a fatal error in this case.



ease of pairing procedure, and ease of re-connecting. This metric is the least objective to measure, but can have a large impact on sales.

See the Appendix for examples of user survey questions related to measuring User Satisfaction.



5 Application of Metrics

Users of the following metrics will benefit from the example task flows given below as they illustrate the application of the metrics.

5.1 JUST WORKS TASK FLOW¹¹

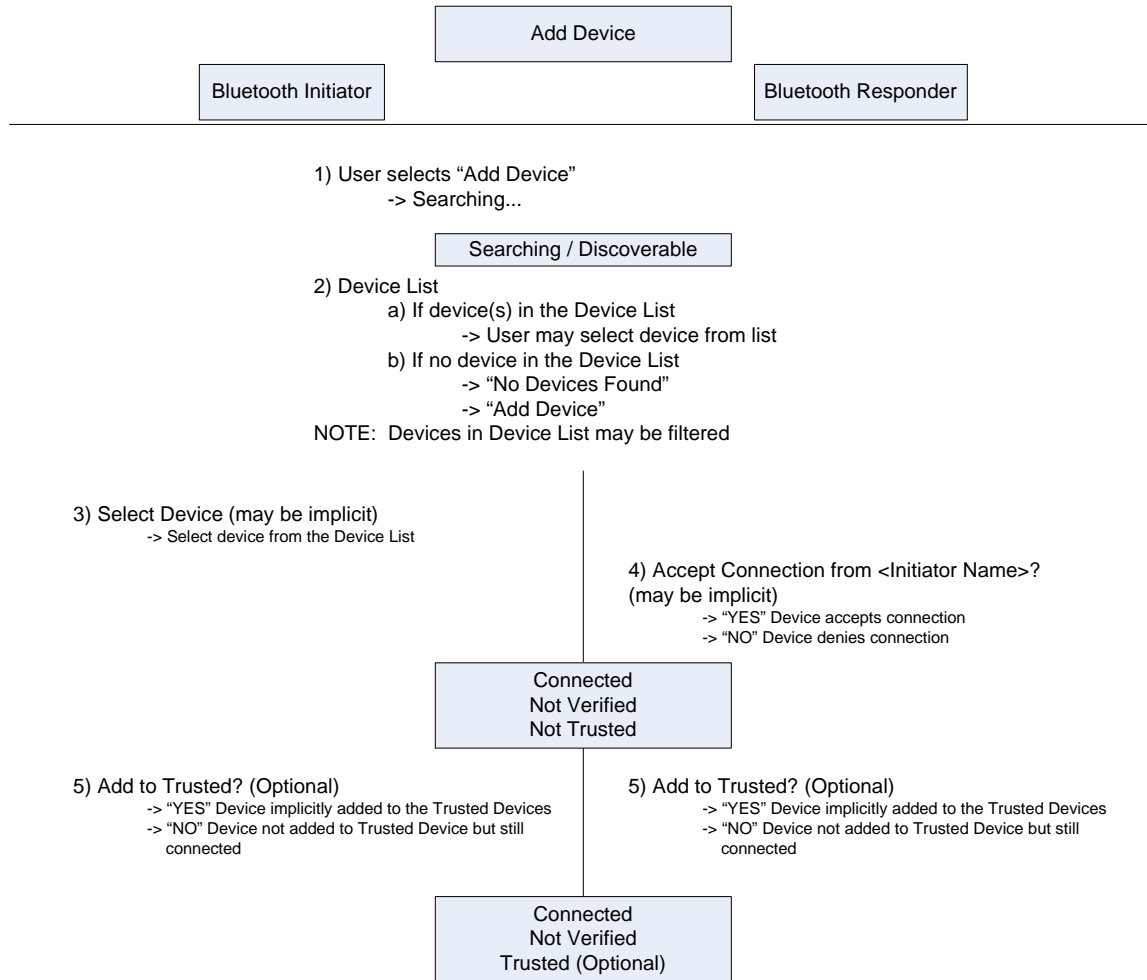


Figure 1: Just Works Task Flow

The following are examples of the Usability Metrics applied to the Just Works Task Flow as shown above. Except for the *Bluetooth* Secure Simple Pairing Task Time, defined by noting starting and ending steps, this is not an exhaustive list of examples.

5.1.1 Bluetooth Secure Simple Pairing Task Time

Begins with Step 1 and ends after Step 5.

5.1.2 Error Count Examples

In Step 2, Responder's device was not in the Device List due to user error.

¹¹ See Usability Expert Group - Secure Simple Pairing UI Flows document for task flow definitions
August 9, 2007



In Step 3, Initiator selects wrong device from the Device List.

In Step 4, Responder selects No on accept and should have selected Yes.

In Step 5, Responder selects No when they intended to select Yes.

In Step 5, Responder selects Yes when they intended to select No.

5.1.3 Fatal Error(s) Examples

In Step 3, Initiator selects wrong device from the list.

5.1.4 Task Incompletion Rate Examples

As an example, if 50 participants attempt the *Bluetooth* Secure Simple Pairing process and 45 complete the task, there is a 10% Task Incompletion Rate. Each device should create its own baseline to determine appropriate success levels.

5.1.5 Intuitiveness Examples

See the Appendix for examples.

5.1.6 User Satisfaction Examples

See the Appendix for examples.



5.2 CONNECT TASK FLOW¹²

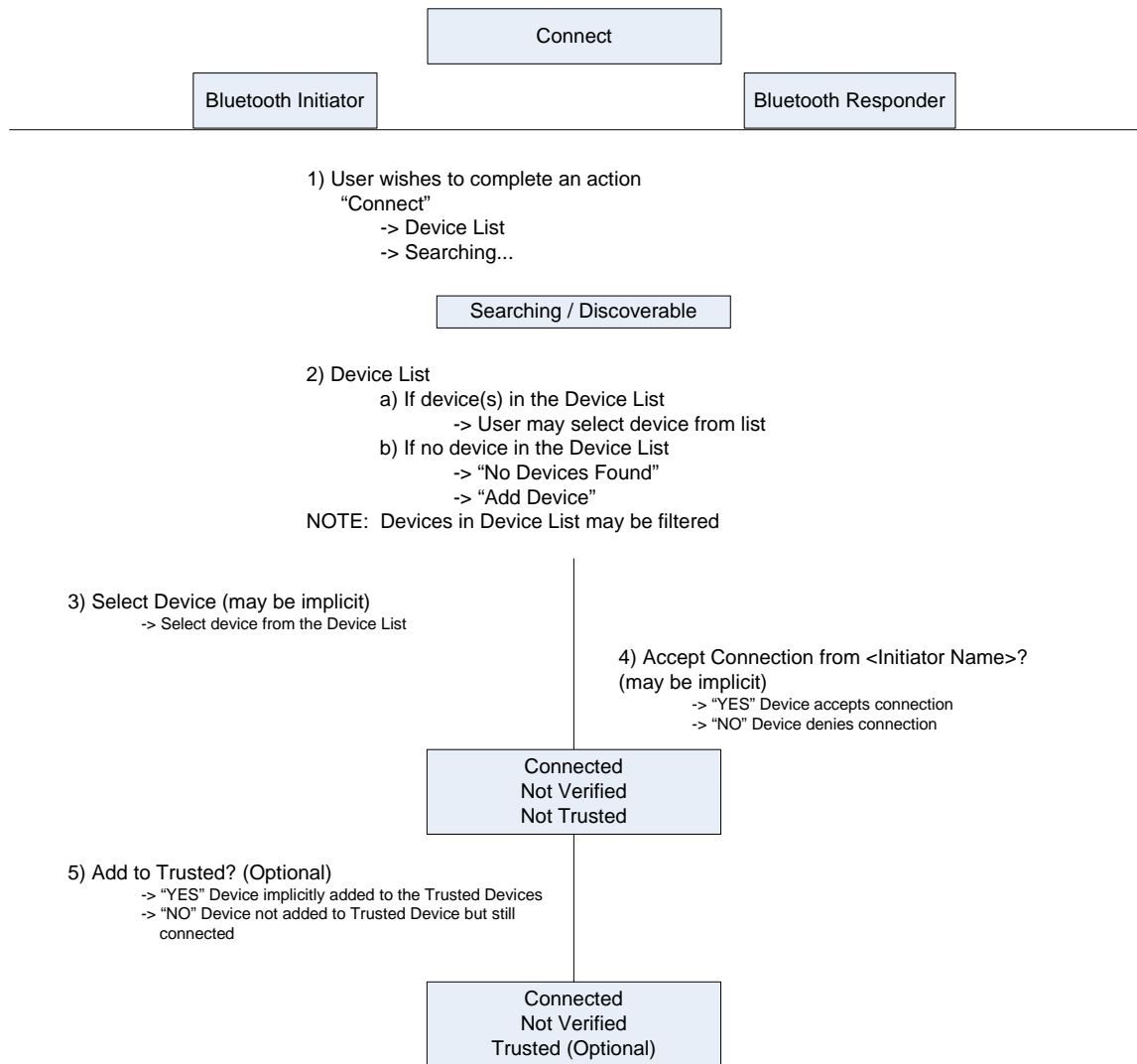


Figure 2: Connect Task Flow

The following are examples of the Usability Metrics applied to the Connect Task Flow as shown above. Except for the Overall Task Time, defined by noting starting and ending steps, this is not an exhaustive list of examples.

5.2.1 Overall Task Time

Overall Task Time = User Intent + *Bluetooth* Secure Simple Pairing.

Begins with Step 1 and ends at the completion of Step 5.

5.2.2 Error Count Examples

In Step 1, Responder's device was not in the Device List due to user error.

In Step 1, Initiator selects wrong device from the Device List.

In Step 2, Responder's device was not in the Device List due to user error.

In Step 3, Initiator selects wrong device from the Device List.

¹² See Usability Expert Group - Secure Simple Pairing UI Flows document for task flow definitions
August 9, 2007



In Step 4, Responder selects No on accept and should have selected Yes.

In Step 5, Responder selects No when they intended to select Yes.

In Step 5, Responder selects Yes when they intended to select No.

5.2.3 Fatal Error(s) Examples

In Step 1, Initiator selects wrong device from the list.

In Step 3, Initiator selects wrong device from the list.

5.2.4 Task Incompletion Rate Examples

As an example, if 50 participants attempt to connect the device and 45 complete the task, there is a 10% Task Incompletion Rate. Each device should create its own baseline to determine appropriate success levels.

5.2.5 Intuitiveness Example

See the Appendix for examples.

5.2.6 User Satisfaction Examples

See the Appendix for examples.



5.3 SEND / RECEIVE TASK FLOW¹³

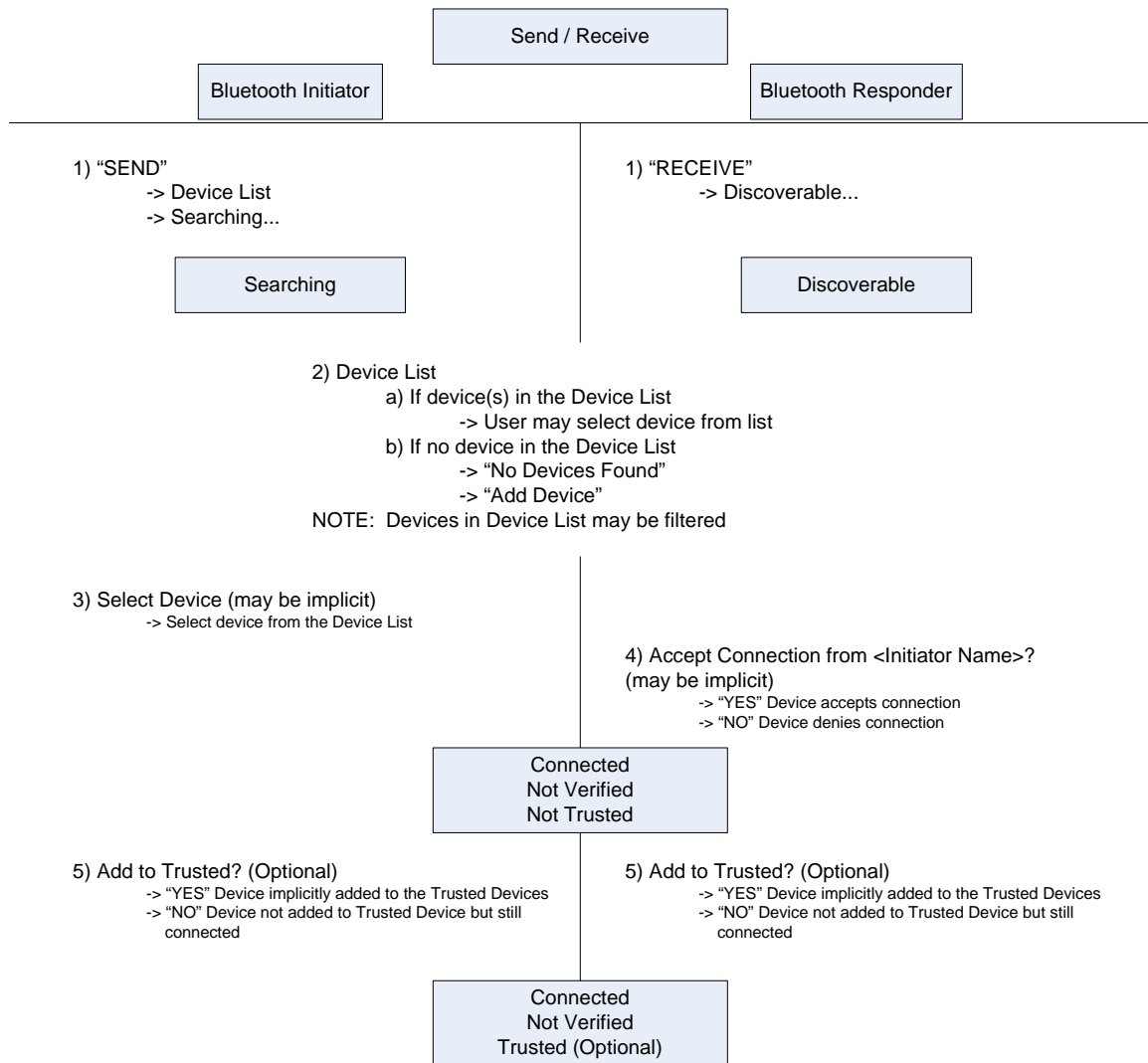


Figure 3: Send/Receive Task Flow

The following are examples of the Usability Metrics applied to the Send / Receive Task Flow as shown above. Except for the Overall Task Time, defined by noting starting and ending steps, this is not an exhaustive list of examples.

5.3.1 Overall Task Time

Overall Task Time = User Intent + *Bluetooth* Secure Simple Pairing.

Begins with Step 1 and ends at the completion of Step 5.

5.3.2 Error Count Examples

In Step 1, Responder's device was not in the Device List due to user error.

In Step 1, Initiator selects wrong device from the Device List.

In Step 1, Responder does not receive file due to user error.

In Step 1, Initiator does not send file due to user error.

¹³ See Usability Expert Group - Secure Simple Pairing UI Flows document for task flow definitions



In Step 2, Responder's device was not in the Device List due to user error.

In Step 3, Initiator selects wrong device from the Device List.

In Step 4, Responder selects No on accept and should have selected Yes.

In Step 5, Responder selects No when they intended to select Yes.

In Step 5, Responder selects Yes when they intended to select No.

5.3.3 Fatal Error(s) Examples

In Step 1, Initiator selects wrong device from the list.

In Step 1, Responder does not receive file due to user error.

In Step 1, Initiator does not send file due to user error.

In Step 3, Initiator selects wrong device from the list.

5.3.4 Task Incompletion Rate Examples

As an example, if 50 participants attempt to send or receive a file and 45 complete the task, there is a 10% Task Incompletion Rate. Each device should create its own baseline to determine appropriate success levels.

5.3.5 Intuitiveness Examples

See the Appendix for examples.

5.3.6 User Satisfaction Examples

See the Appendix for examples.



Appendix

Example *Bluetooth* Secure Simple Pairing Survey

#	Question	Strongly Agree	Somewhat Agree	Neither Agree nor Disagree	Somewhat Disagree	Strongly Disagree
<i>Intuitiveness</i>						
1	I thought the <i>Bluetooth</i> Secure Simple Pairing process was intuitive.					
2	I agree I did not need to consult a manual to complete the <i>Bluetooth</i> Secure Simple Paring process.					
3	The <i>Bluetooth</i> Secure Simple Pairing process worked how I expected.					
4	I did not need any additional assistance to complete the <i>Bluetooth</i> Secure Simple Pairing process.					
5	I thought the <i>Bluetooth</i> Secure Simple Pairing user interface was clear and easy to follow.					
6	I would imagine that most people would be able to pair these devices quickly.					
7	I think I would need the support of a technical person to be able to pair these devices.					
<i>User Satisfaction</i>						
8	I was satisfied with the <i>Bluetooth</i> Secure Simple Pairing process.					
9	I thought the <i>Bluetooth</i> Secure Simple Pairing process was easy to complete.					
10	I thought the <i>Bluetooth</i> Secure Simple Pairing instructions that were provided to me with the product were inadequate.					
11	These devices worked just the way I expected them to work.					
12	I thought the user interface options and their functions were inconsistent.					
13	I felt very confident pairing these devices.					
14	I thought the <i>Bluetooth</i> Secure Simple Pairing process to be very complicated.					
15	I thought the <i>Bluetooth</i> Secure Simple Pairing process unnecessarily complex.					
16	Based upon my experience with <i>Bluetooth</i> Secure Simple Pairing I would recommend <i>Bluetooth</i> devices to my friends.					
17	Based upon my experience with <i>Bluetooth</i> Secure Simple Pairing I would purchase <i>Bluetooth</i> devices.					