Quality of PDA-Based M-Learning Information Systems

Ruti Gafni
The Open University of Israel
rutiga@openu.ac.il

Abstract
PDA-based information systems are suitable for m-learning, enabling ubiquitous and idle-time utilization. However, the use of small devices and mobile systems requires special quality attention in order to get the desired benefits. This article examines the quality problems resulting from the characteristics of these systems, proposes some metrics to measure their quality and some practical implications for the designers and developers of PDA-based systems, including technical, implementation and content issues.

Keywords: personal digital assistant (PDA), m-learning, information systems usability, mobile, quality.

Introduction
Mobile learning (m-learning) enables learning, independently of place and time, ubiquitous, through wireless networks and mobile devices, such as personal digital assistants (PDA), cellular phones, mp-3 devices and others (Alexander, 2004). Using PDA, the system can be activated online, while connected to the network, or offline, in case the information was previously loaded to the device memory, and then used when disconnected from the network.

Use of mobile systems increased, and will further grow (Lee, 2003; Lau, 2006), due to the emergence of third generation wireless networks and since technology, devices and infrastructures matured. This has lead to development of diverse kinds of systems, including m-learning, which can be used by a variety of users, for different tasks. Among others, students traveling by bus or train can read course summaries, perform online quizzes or assessments, and send them to the relevant instructor; company workers can use these systems for specific training, using any free time, wherever workers are situated; travelers can avoid carrying travel books when visiting old towns, and receive the desired information using their mobile device.

M-learning systems can create benefits for organizations and users; e.g., better idle time utilization, productivity enhancement, flexibility of procedures, etc., which together emphasize competitive strategy, lower operation costs, and improve processes. However, these benefits can be nullified by using sub-quality information systems (Terho, 2002). Mobile-wireless information systems face new kinds of problems, different from traditional and internet systems, which have been largely researched (Kan, 2002; Calero et al., 2004; Covella & Olsina, 2006; Pandian, 2004). These problems include small screens, keyboards and memory, narrow bands, lack of coverage and diversity of users and devices. M-learning systems are affected by these special problems; thus, they require an enhanced quality definition. The purpose of this study is to find metrics to quantify the quality of such information systems, in order to enable the evaluation, comparison and analysis of m-learning information systems quality.

This article deals with PDA-based m-learning information systems with offline activation. Relevant metrics are defined and validated using a PDA-based experiment.

Proceedings of the Chais conference on instructional technologies research 2009: Learning in the technological era
Y. Eshet-Alkalai, A. Caspi, S. Eden, N. Geri, Y. Yair (Eds.), Raanana: The Open University of Israel

Information Systems Quality Measuring

Measuring creates a quantitative description which allows behavior comprehension and enables selection of tools and techniques to control and improve processes, products and resources. Information systems quality cannot be measured only by software faults absence; it must be broader, including characteristics to cover all aspects, life-cycle phases and viewpoints. Several hierarchical models for quality definition were proposed (McCall, Richards, & Walters, 1977; Boehm, Brown, Kaspar, Lipow, & McCleod, 1978). The most recent is ISO/IEC 9126 standard (ISO/IEC9126-1, 2001), which decomposes quality into seven characteristics, further divided into sub-characteristics (Figure 1). This standard defines, for each sub-characteristic, internal metrics to be measured without having to operate the system and external metrics to be measured while testing or executing the system. Gafni (2008) proposed to extend the metrics of the standard to cope with the problems of the mobile-wireless information systems.

Quality Issues in PDA-based M-learning Systems

Mobile-wireless information systems face some unique problems originating from the mobile devices and from the characteristics of the network. When used offline, these systems suffer especially from problems derived from the mobile devices:

- Technical limitations: small memories (the amount of information stored on it cannot be large), short battery life (the continuous period of work cannot be long) and limited calculation and computation capabilities.
- Wide variety of devices, possessing different characteristics, which the application must be adaptable to all of them (Brady, 2004).
- Uncomfortable use of devices: tiny screens, which restrict the amount of data displayed, and small keyboards that are difficult to operate.
- Security, privacy and confidentiality problems (Di Pietro & Mancini, 2003; Herzberg, 2003), can arise when lost, due to possible unauthorized access to sensitive data.

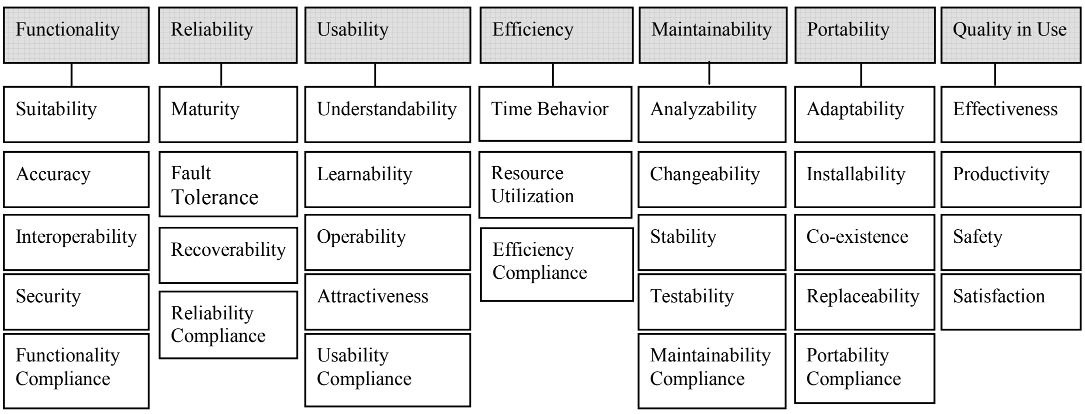

Figure 1 – ISO/IEC 9126 quality characteristics and sub-characteristics

These problems intensify when the m-learning system is targeted for a wide audience, where the users are faceless and there is neither user training nor implementation, but they are significant also in closed organizational information systems (firms, schools, universities, etc.).

Using PDA-based m-learning systems, instead of smaller devices like cell phones, can be beneficial because the devices are larger, they include Qwerty style keyboards, or stylus which allows the user to interact with the system, by touching the screen, and data can be gathered in memory to be further sent to the remote system. Even though, the PDA devices do not solve the problems listed above, and demand precise considerations and measurement of the system quality, significantly influenced by all these features, using new defined metrics. This article deals with measuring quality of the system enabling learning using PDA devices, and not the quality of the learning itself. It covers the system quality level, identified by Kim and Ong (2005) as one of the three success factors of m-learning from the learners' perspective.

Suitability means that the application fits the needs to fulfill a certain task without overwhelming the user. This is important in m-learning systems, because of the use of small devices. For this reason the information must be delivered in small and concise amounts (Venkatesh & Ramesh, 2002; Parsons & Ryu, 2006).

Operability is affected mainly by the mobile device attributes which restrict output and input interaction possibilities. Output capabilities of devices are determined by their screen, which limits the amount of information for simultaneous display. The input restrictions are due to size constraints of mobile devices and due to keyboard capabilities.

Since mobile applications may be used while driving or walking, tasks which need the user attention, the application manipulation must be simple and intuitive (Terrenghi et al., 2005), enforcing safety. Moreover, the noisy surroundings, may distract the user and cause input errors, inaccuracy and slowness.

Portability is very important in m-learning systems which are targeted to wide audiences, because of the large diversity of devices held by users and the lack of control over these devices. The application must adapt itself to the device features, both during installation and operation, and according to the user's preferences (Goh & Kinshuk, 2006). The application installation on the device must be invisible to laypersons.

Quality Metrics for PDA-based M-learning Systems

M-learning systems must be measured on the basis of traditional systems metrics, e.g. ease of maintainability, minimum complexity, lack of faults, etc., which are fundamental metrics for all information systems. Furthermore, when used online, they must be measured according to internet systems metrics, e.g. no broken links, ease of navigation, etc. In addition, they need to be measured with special targeted mobile-wireless metrics (Spriestersbach & Springer, 2004).

The definition of new metrics to quantify the quality of m-learning systems is based on Gafni's (2008) research process, which consists of several phases:

(1) detection of quality problems and risks that outcome from the architectures and protocols of mobile-wireless information systems (Asunmaa et al., 2002; Green, 2003; Huber, 2004; Tarasewich, 2003; Varshney & Vetter, 2002; Vaughan-Nichols, 2004).
(2) Choice of ISO/IEC 9126 quality characteristics affected by mobility and wirelessness.
(3) Identification of objects (clustered into four entities: device, application, architecture and user), which were decomposed into measurable attributes (ISO/IEC 15939, 2002) and assigned a unit measure and a scale, according to its meaning (Kitchenham, 1995).
(4) Definition of metrics, methods for the measuring process with specific formulas, (ISO/IEC 15939, 2002) to allow objective measurement of mobile-wireless information systems quality.
(5) Metrics theoretical and empirical validation.

Table 1 presents part of defined and evaluated metrics for measuring the quality characteristics of a PDA-based m-learning system. Table 2 maps these metrics according to the different kind of problems which they measure.

Table 1 – PDA-based m-learning systems quality metrics

Metric Name	**Metric purpose**	**Method of calculation**
Response time to get information from cache (local memory)	Measures the response time to get information stored in the device's memory. The efficiency will be greater when the response time is smaller.	Operate same task number of times. First, the data will be received from server and stored in device's memory. Next, the data will be retrieved locally. The time to get data as output is measured, starting from second iteration.
Device memory cleanup after transaction	Measures the degree of cleaned-up memory after completing a task. The small resources of memory in the devices have to be carefully handled to avoid decrease of performance.	Check free memory size before and after performing task. Check number of times to see if size decreases.
Display load	Measures the burden degree of the displays. Because of the screen's size, the understandability is lower when the display is overloaded.	Calculate ratio between display size and device screen size, for each display in system. Distribute ratios to inversed weighted categories, and calculate weighted average.
Messages conciseness	Measures the conciseness of operational and error messages. This save place in memory and the messages are easy to read from screen.	Calculate ratio between message size and maximum lines in screen. Distribute ratios to inversed weighted categories, and calculate weighted average.
Ease of input entering	Measures the ratio of easy to fill input fields, like usage of default values, list of values or self-completing fields.	Count number of closed input fields, default value fields and self-completing fields. Calculate ratio between counted fields and total of input fields.
Ease of output use	Measures the suitability between length of outputs and device screen size.	Calculate average displays for message. Count number of outputs needing leafing, by checking ratio between number of lines in output message and number of lines in screen.
Ease of use – displays per task	Measures the number of screens involved in one task. The effectiveness and usability is greater when the system is easier to operate.	Calculate number of iterations between user and system, till end of task.
Secure messages and information on device	Measures the usage degree of security mechanisms, like identification, authorization and	Check if the system uses mechanisms to secure information. Categorize the degree of usage.

Metric Name	Metric purpose	Method of calculation
	confidentiality.	
Use of user profile	Measures the extent in which the system relays on user profile to fit the inputs and outputs to the user and device.	Check if the system uses mechanisms to fit to users profile. Categorize the degree of usage.
Installations success	Measures the degree of successful installation in the different planned devices.	Calculate number of successful installations according to devices types and total installations.
Ease of installation	Measures the duration time and ease of installation.	Calculate time of system installation on devices.

Table 2 – Metrics mapped to mobile-wireless problems

Source of problems:	Network					Device						Mobility			
Type of problem / Metrics	Narrow band	Connection stability	Security	Diverse standards	High costs	Variety of devices	Low memory	Tiny screens	Limited computational abilities	Small keyboard	Loss and information security	User adapted information	Security and privacy	Disturbances and disconnections	Users trust
Response time to get information from cache	X		X		X		X		X					X	X
Device memory cleanup after transaction							X								
Display load								X							
Messages conciseness	X				X		X	X							
Ease of input entering										X					
Ease of output use	X	X	X		X									X	X
Ease of use – displays per task	X	X			X			X	X					X	X
Secure messages and information on device											X	X	X		X
Use of user profile	X				X	X						X			
Installations success						X									
Ease of installation						X									

The "PDA" Experiment

Each metric defined by the research was validated theoretically, mathematically, and empirically. The aim of the validation was to prove that the metrics behave in a consistent and logical mode. Thus, it is possible to rely on them to quantify the quality of mobile-wireless information systems, for example, by showing that the value of the metric grows when quality increases, and vice versa.

This research was part of a broader research, to define new quality metrics, in which different kind of mobile-wireless applications were inspected. Some of the metrics were compared between different experiments. This paper focuses on the PDA application, and the validation in the "PDA" experiment alone.

The metrics were validated empirically by performing the "PDA" experiment, based on a system developed by Dooblo, a commercial company. The experimental system contains a set of questionnaires, which allow conducting surveys, using a PDA device based on Windows Mobile CE operation system. The system contains five different surveys, "Supermarket", "Purchase", "Participation", "Movie" and "Flight". Each survey includes 6-10 displays (samples in Figure 2). These questionnaires include different kinds of questions, like: open questions, selecting an answer from a number of possibilities, etc., similar to performing quizzes in a PDA-based m-learning system. The questionnaires were loaded to the PDA device's memory in advance. The surveys were performed, accumulating the data on the device, and further downloaded to a server for elaboration.

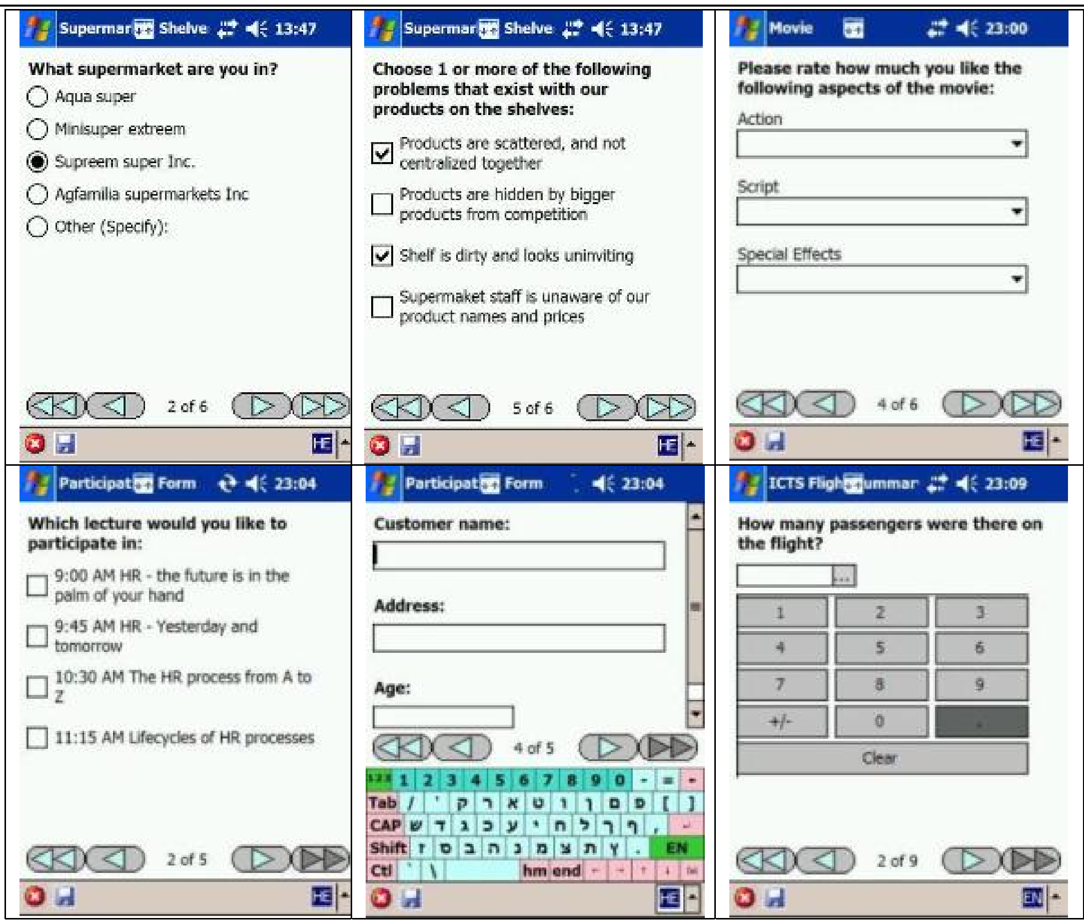

Figure 2 – Samples of "PDA" experiment displays

This is a realistic experiment, using real tasks and environment (Sjoberg et al., 2002), and the observation was done as a "black box", with no possibility to see or interfere with the code of the system. Each survey was performed thirty times, and the data needed for the different metrics was collected.

The theoretical proof and the empirical experiments, summarized in Table 3, successfully validated the new metrics defined in the research.

Table 3 – Results of metric validations in PDA experiment

Metric Name	**Result of validation**
Response time to get information from cache	For displays with longer response time a lower value was received for the metric.
Device memory cleanup after transaction	In the cases where the memory was better cleaned after completing the task, the metric value was higher, as shown in Figure 3.
Display load	For less loaded displays the metric value received was higher, i.e., the displays of the "Flight" survey which are the less loaded, got a metric value of 1, the highest possible. (Figure 4).
Messages conciseness	Surveys in which the messages were shorter in proportion with display, got a higher metric value.
Ease of input entering	For surveys where the ratio of "closed" questions was higher, the value received was higher, as shown in Figure 5.
Ease of output use	When the number of leafing needed to see all the output is greater, the value of the metric received was smaller.
Ease of use – displays per task	When the number of iterations needed to finish a task is greater, the value of the metric received was smaller.
Secure messages and information on device	Validated through other experiments.
Use of user profile	Validated through other experiments.
Installations success	Validated through other experiments.
Ease of installation	Validated through other experiments.

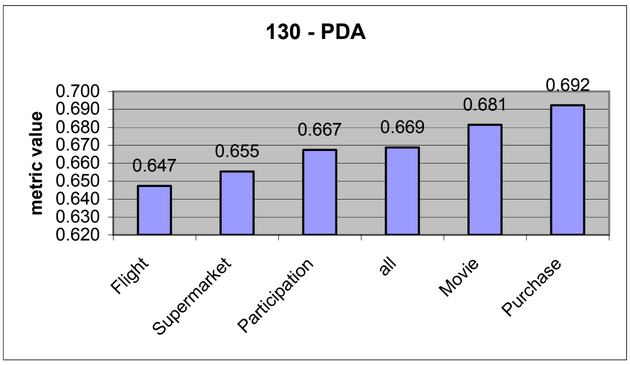

Figure 3 – Validation "Memory Cleanup" metric in "PDA" experiment

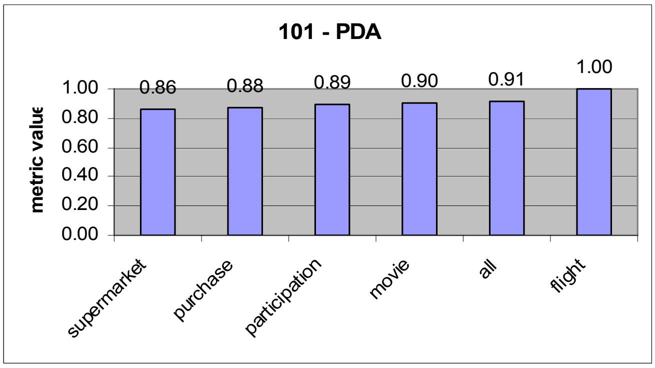

Figure 4 – Validation of "Display load" metric in "PDA" experiment

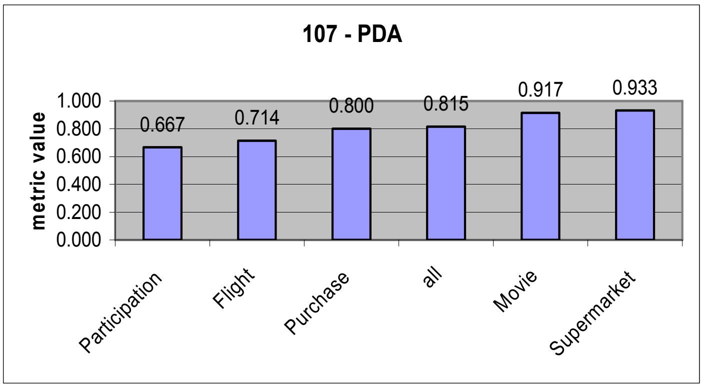

Figure 5 – Validation "Ease of input entering" metric in "PDA" experiment

Practical Implications

Some practical implications can be deducted from this research, including technical, implementation and content issues, to assist the developers of PDA-based m-learning systems.

The technical developers must be conscious of the different kinds of devices which will be used, including differences in memory size, screen size and attributes, to design an application suitable to all. This can be accomplished by using user profiles mechanisms. Applications must be aware of the small memory and low processing resources restrictions and optimize their utilization.

The content developers must be aware of the restrictions in size and usability, limiting input to minimal required data, determining pre-defined options and automatic filled-in fields, using location aware information when possible, instead of user's input. They need to consider output length, by defining concise messages and adapting the content of the application to screen size, taking into consideration that the operation of leafing through multiple pages in a mobile device decreases the usability.

Future Enhancements

The purpose of the "PDA" experiment was to validate the definition of the metrics, and this is the reason the small sample was enough. The experiment can be enhanced with an extended sample. In this experiment, only one PDA application was examined, in one kind of device. The experiment can be broadened by enlarging the number of kinds of devices and the number of applications. Moreover, it can be performed again, with a real m-learning information system, instead of the survey system.

Summary

This article describes a methodology to enable quality measuring of m-learning systems which allow interactive learning "any place, any time" using PDA devices. PDA-based m-learning systems are a sub-field of mobile-wireless information systems, which face unique problems and challenges that affect quality attributes. These attributes were analyzed and accurate metrics were defined. These metrics enable objective quality evaluation and comparison of mobile-wireless information systems.

These metrics are useful when the quality of a mobile-wireless information system must be analyzed and quantified, for example when comparing two proposed systems, or when a system has to be developed or bought. When the metrics are used to compare systems, the higher the metric value, the higher the system's quality. However, when only one system has to be measured, the metrics need an external value to compare to. These values can be defined in advance according to the requirements of the system. Moreover, practical implications for the development of PDA-based m-learning systems were defined based on insights gained from the metrics definition and methodology.

References

Alexander, B. (2004). Going nomadic: Mobile learning in higher education. *EDUCAUSE Review*, *39*(5), 28–35.

Asunmaa, P., Inkinen, S., Nykänen, P., Päivärinta, S., Sormunen, T., & Suoknuuti, M. (2002). Introduction to mobile internet technical architecture. *Wireless Personal Communications*, *22*, 253–259.

Basili, V. R., & Rombach, H. D. (1988). The TAME project: Towards improvement-oriented software environments. *IEEE Transactions on Software Engineering*, *14*(6), 758-773.

Boehm, B. W.; Brown, J. R.; Kaspar, J. R., Lipow, M.; McCleod, G. (1978). *Characteristics of software quality*. Amsterdam: North Holland.

Brady, A., Conlan, O. & Wade, V. (2004). Dynamic composition and personalization of PDA-based eLearning – Personalized m-learning. In G. Richards (Ed.), *Proceedings of World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education 2004*, *Chesapeake, VA: AACE,* 234-242.

Calero, C., Ruiz, J., & Piattini, M. (2004). A Web metrics survey using WQM. *Proceedings ICWE 2004, LNCS 3140, Springer-Verlag Heidelberg* , 147–160.

Covella, G., & Olsina, L. (2006). Assessing quality in use in a consistent way. Proceedings of the 6th International Conference on Web Engineering, ICWE'06, Palo Alto, California, USA, 1-8.

Di Pietro, R., & Mancini, L. V. (2003). Security and privacy issues of handheld and wearable wireless devices. *Communications of the ACM*, *46* (9), 74-79.

Green, R. (2003). Wap 2.0: what is it? Wap 2.0 is coming, sooner than you might think, and it's going to change everything. *Wireless Business and Technology*, *3*(6), 34-38.

Gafni, R. (2008). Framework for quality metrics in mobile-wireless information systems. *Interdisciplinary Journal of Information, Knowledge, and Management*, *3*, 23-38.

Goh T., Kinshuk (2006). Getting ready For mobile learning – Adaptation perspective. *Journal of Educational Multimedia and Hypermedia, 15* (2), 175-198.

Herzberg, A. (2003). Payments and banking with mobile personal devices. *Communications of the ACM*, *46*(5), 53-58.

Huber, J. F. (2004). Mobile next-generation networks. *IEEE Multimedia*, *11*(1), 72-83.

ISO/IEC 15939, (2002). Software Engineering – Software Measurement Process. Geneva, Switzerland: International Organization for Standardization.

ISO/IEC 9126, (2001). Software engineering – Product quality – Part 1: Quality model. Geneva, Switzerland: International Organization for Standardization.

ISO/IEC 9126, (2004). Software engineering – Product quality – Part 4: Quality in use metrics. Geneva, Switzerland: International Organization for Standardization.

Lau, J. (2006). The state of European enterprise mobility in 2006. *Forrester Research*.

Lee, P. (2003). Mobile data comes of age: wireless is big, we know that, but how big? *Wireless Business and Technology, 3*(9), 16-18.

Kan, S. (2002). Metrics and models in software quality engineering (2nd Edition), Addison-Wesley.

Kim, G. & Ong, S. M. (2005) An exploratory study of factors influencing m-learning success. *The Journal of Computer Information Systems*, Oct 1, 2005.

Kitchenham, B., Pfleeger, S. L., & Fenton, N. (1995). Towards a framework for software measurement validation. *IEEE Transactions on Software Engineerin*g, *21*(12), 929-944.

Malladi, R., & Agrawal, D. (2002). Current and future applications of mobile and wireless networks. *Communications of the ACM*, *45*(10), 144-146.

McCall, J.A., Richards, P. K., & Walters, G.F. (1977). Factors in software quality. RADC TR-77-369, Vols I,II,III', US Rome Air Development Center Reports NTIS AD/A-049-014, 015, 055.

Pandian, C. R. (2004). Software metrics: A guide to planning, analysis, and application, Auerbach Publications.

Parsons, D., & Ryu, H. (2006). A Framework for assessing the quality of mobile learning. Proceedings of the 11th International Conference for Process Improvement, Research and Education (INSPIRE), Southampton Solent University, UK.

Schiller, J. (2000). *Mobile communications*. Addison-Wesley.

Sjøberg, D.I.K., Anda, B., Arisholm, E., Dybå, T., Jørgensen, M., Karahasanovic, A., Koren, E. F.; Vokác, M. (2002). Conducting realistic experiments in software engineering. *Proceedings of the 2002 International Symposium on Empirical Software Engineering (ISESE'02)*, 17-26.

Stafford, T. F., & Gillenson, M. L. (2003). Mobile commerce: What it is and what it could be. *Communications of the ACM*, *46*(12), 33-34.

Spriestersbach, A., & Springer, T. (2004). Quality attributes in mobile web application development. *LNCS 3009, Proceedings of PROFES, Berlin*: Springer-Verlag, 120-130.

Tarasewich, P. (2003). Designing mobile commerce applications. *Communications of the ACM*, *46*(12), 57-60.

Terho, M. (2002). Mobile web services and software quality. *LNCS 2349, Proceedings of ESCQ, Berlin*: Springer-Verlag, 2-6.

Terrenghi, L., Kronen, M., & Valle, C. (2005). Usability requirements for mobile service scenarios. *Proceeding of HCI International Conference, Las Vegas, USA*, 1-10.

Varshney, U., & Vetter, R. (2002). Mobile commerce: framework, applications and networking support. *Mobile Networks and Applications*, *7*, 185–198.

Vaughan-Nichols, S. J. (2004). Wireless middleware: Glue for the mobile infrastructure. *IEEE Computer*, *37*(5), 18- 20.

Venkatesh, V., Ramesh, V. (2002). Usability of web and wireless sites: Extending the applicability of the Microsoft usability guidelines instrument. *Information Systems Technical Reports and Working Paper, TR134-1*; Retrieved Nov. 20, 2006 from http://www.kelley.iu.edu/ardennis/wp/tr134-1.doc