Interacting with Computers 22 (2010) 75–87

Contents lists available at ScienceDirect

Interacting with Computers

journal homepage: www.elsevier.com/locate/intcom

The usability inspection performance of work-domain experts: An empirical study

Asbjørn Følstad [a,*], Bente C.D. Anda [b], Dag I.K. Sjøberg [b]

[a] Department of Cooperative and Trusted Systems, SINTEF, Oslo, Norway
[b] Department of Informatics, University of Oslo, Norway

ARTICLE INFO

Article history:
Received 31 January 2008
Received in revised form 20 August 2009
Accepted 1 September 2009
Available online 6 September 2009

Keywords:
Usability inspection
Work-domain expert
Validity
Thoroughness
Group-based expert walkthrough

ABSTRACT

It is a challenge for usability experts to perform usability inspections of interactive systems that are tailored to work-domains of which these experts have little knowledge. To counter this, usability inspections with work-domain experts have been explored, but little empirical research has been reported on these experts' performance as evaluators. The present study compared the performance of work-domain experts and usability experts with respect to validity and thoroughness. The work-domain experts were characterized by high computer experience and low system experience. The usability experts were recruited from different ICT companies. The usability inspection method applied was group-based expert walkthrough; a method particularly developed to support non-usability experts as evaluators. The criterion for performance comparison was established through user tests. Fifteen work-domain experts and 12 usability experts participated in the study. The work-domain experts generated equally valid but less thorough usability inspection results than did the usability experts. This finding implies that work-domain experts may be used as evaluators in usability inspections without compromising validity. Moreover, the usability inspection performance of nominal groups of evaluators was explored. It was found that nominal groups of work-domain experts produced results of similar quality as did nominal groups of usability experts, given that group size is disregarded. This finding may be used as basis for hypotheses in future studies on the usability inspection performance of nominal groups of work-domain experts.

1. Introduction

Since the late 1980s, usability inspection, or analytical evaluation, has been promoted as a cost-effective alternative to empirical evaluation (such as user tests) for identifying usability problems in interactive systems (Cockton, Lavery, & Woolrych, 2003). Today, usability inspection methods are widely used; in particular heuristic evaluation, and to some degree cognitive walkthrough (Bark, Følstad, & Gulliksen, 2005; Venturi and Troost, 2004). These methods are typically applied by usability experts (Virzi, 1997).

The usability of an interactive system depends on the context of the system (ISO, 1998). Therefore, usability evaluations need to be conducted in consideration with the context (Bødker and Madsen, 1998). However, usability inspection methods such as heuristic evaluation and cognitive walkthrough do not include sufficient context knowledge resources, which represents a "major source of inspection risk" (Cockton, Lavery, & Woolrych, 2008, p. 1179). In consequence, it is a challenge for usability experts to perform usability inspections of interactive systems that are tailored to work-domains of which these experts have little knowledge. To acquire adequate knowledge of the work-domain context, usability experts may need to receive extensive introduction or training, or conduct context research. Such activities may represent a prohibitive overhead.

Motivated by the challenge of ensuring that sufficient knowledge of the system's context is available, alternative approaches to usability inspections have been explored. One approach has been to involve work-domain experts as evaluators in the usability inspection. Work-domain experts are "(a) potential end users with direct experience from the work-domain or (b) persons with extensive secondary knowledge of the work-domain" (Følstad, 2007b, p. 218). Usability inspection methods particularly developed to include work-domain experts as evaluators have existed since the pluralistic walkthrough of Bias (1994), with the participatory heuristic walkthrough (Muller, 1998) and group-based expert walkthrough (Følstad, 2007a) as more recent additions.

When work-domain experts are employed as evaluators in usability inspections, required knowledge of the work-domain context is ensured. However, the inspection performance will be

Abbreviations: HCI, human–computer interaction; ICT, information and communication technology; WDE, work-domain experts; UE, usability experts.

* Corresponding author. Address: SINTEF ICT, P.O. Box 124, Blindern, 0314 Oslo, Norway. Tel.: +47 22067515; fax: +47 22067350.
E-mail addresses: asf@sintef.no (A. Følstad), bentea@ifi.uio.no (B.C.D. Anda), dagsj@ifi.uio.no (D.I.K. Sjøberg).

0953-5438/$ - see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.intcom.2009.09.001

affected if the usability intuitions of the work-domain experts are not sufficiently strong. In particular, unsound usability judgements may decrease the validity of the usability inspection. This is potentially critical, given the risk of false positives which seem to be associated with usability inspections (e.g. Cockton et al., 2008).

Little research has been conducted on the performance of work-domain experts' as evaluators in usability inspections, with the consequence that the human–computer interaction (HCI) community has little knowledge in this regard. This is particularly problematic with respect to industrial application of usability inspections involving work-domain experts as evaluators. This article reports on a study that compared the performance of work-domain experts and usability experts as evaluators in usability inspections conducted as group-based expert walkthroughs. The main performance measures of the study were the validity and thoroughness of the usability inspection results relative to the results of empirical user testing.

2. Previous work

2.1. Work-domain experts in human–computer interaction

The importance of involving work-domain experts in the development of interactive system is highlighted by a number of HCI researchers. In particular within the participatory design tradition, the value of mutual and reciprocal meetings of users, as experts on their work-domain, and developers has been emphasized (Muller, 2008). Kensing and Munk-Madsen (1993) argued that such involvement of the users is critical in order to "build up the knowledge required for developing and using a new system" (p. 78). To explain the value of the knowledge possessed by work-domain experts, Kensing and Munk-Madsen established six categories to describe knowledge required for adequate understanding between users and developers. Two of these knowledge categories included knowledge typically possessed by the work-domain experts alone: (1) concrete experience with the users' present work and (2) knowledge about relevant structures on this work.

To understand the significance of context knowledge possessed by work-domain experts in the development projects, it is useful to consider Nielsen's (1993) three dimensions of individual differences between users: users' varying degree of (1) computer experience, (2) system experience, and (3) work-domain knowledge (or task-domain as it is termed by Nielsen). When developing interactive systems for users with higher levels of domain knowledge, it will be more critical to involve work-domain experts as knowledge resources.

Nielsen's three dimensions also help in discriminating between different categories of work-domain experts. Work-domain experts will by necessity be high on the dimension of work-domain knowledge, and, conversely, any user with high work-domain knowledge is reckoned a work-domain expert; this characteristic is what separates work-domain experts from other users. However, work-domain experts may vary on the dimensions of (a) specific system experience and (b) general computer experience. In consequence, it may be useful to classify work-domain experts into four categories given by high/low on these two dimensions.

It should be noted that work-domain experts may vary on a range of other dimensions as well. Nielsen (1993) suggests that important differences between users may be related to age, gender, differences in spatial memory and reasoning abilities, as well as preferred learning style. Recent research by Ling and Salvendi (2009) indicates that the evaluators' cognitive style may affect usability inspection performance. On this background, a model of work-domain experts based on differences in system experience and computer experience alone may seem simplistic. Even so, the model does seem to represent a good starting point for making relevant classifications while keeping with a modest level of complexity.

In the theory of usability evaluation, over the last two decades, there has been a movement toward treating the users as work-domain experts rather than subjects just to be observed. With basis in the participatory design tradition, Ross, Ramage, and Rogers (1995) introduced PETRA (participatory evaluation through redesign and analysis) in which users are regarded as experts that may suggest design changes in interwoven design and evaluation sessions. In the field of CSCW (computer supported cooperative work), Twidale, Randall, and Bentley (1994) presented informal evaluation sessions where the participants were used as reflecting experts, in order to capitalize on their context knowledge.

Other researchers have argued that users may be utilized as work-domain experts in user testing. Boren and Ramey (2000) argued that the participating user should be "cast as an important contributor, work-domain expert, or valued customer" (p. 268), an approach that was later empirically investigated by Krahmer and Ummeln (2004). Frøkjær and Hornbæk (2005) presented a new version of user testing, the cooperative usability testing, where the inclusion of an interpretation phase with discussions between the participant and the test leader enables "test users and evaluators join expertise to understand the usability problems of the application evaluated" (p. 1383). Similarly, Aborg, Sandblad, Gulliksen, and Lif (2003) presented an evaluation method where the evaluator performs observation interviews with users during their ongoing work with the evaluated system.

Different usability inspection methods have been developed in order to include work-domain experts as evaluators. In Bias (1994) pluralistic walkthrough users, developers, and usability experts participate as interacting groups of evaluators. In the field of participatory design, the participatory heuristic evaluation was developed (Muller and McClard, 1995; Muller, 1998). This usability inspection method includes heuristics that target the systems' fit to the user needs and work environment, and allows usability experts and work-domain experts to cooperate in usability inspections.

2.2. The inspection performance of work-domain experts

Little is known about the usability inspection performance of work-domain experts. Muller and McClard (1995) presented a validation study of the participatory heuristic evaluation, but they targeted only the effect of the added usability heuristics without singling out the relative contribution of the different evaluators.

However, some knowledge exists on the performance of inexperienced evaluators; this may be relevant because work-domain experts are typically inexperienced evaluators. Nielsen (1992) found that, on an average, it may take between 10 and 15 inexperienced evaluators to identify the same proportion of usability problems that can be identified by three usability specialists.[1] Similarly, Desurvire, Kondziela, and Atwood (1992) found that a group of non-experts identified less than one third of the total number of usability problems than did a group of usability experts. These studies might indicate that both a lower number of problem predictions and lower thoroughness can be expected in the performance of work-domain experts than in that of the usability experts, even though the recent work of Howarth, Smith-Jackson, and Hartson (2009) indicates that analysis support may improve the performance of novice usability evaluators.

[1] Nielsen (1992) also included a third category, double experts, who are experts in both usability and the application domain. Nielsen's *application domain* (e.g. the domain of voice response telephone menus) should not be confused with *work-domain* as used in this article. Nielsen's term refers to a particular category of applications and not on a particular work-context.