

Kabin İçi Eğlence Sistemi Geliştiriminde bir Kullanılabilirlik Testi Uygulaması

Hilal Karaca¹, Emre Er¹, Adem Çağlar¹, Kadir Herkiloğlu¹, Seçkin Çaylak¹, İ. Berk Işık¹, Burak Yolaçan¹

¹ HAVELSAN A.Ş. Mustafa Kemal Mah. 2120 Cad. No:39 Çankaya, Ankara, Türkiye
{hkaraca, xscaylak, iisik}@bgt.havelsan.com.tr,
{eer, acaglar, kherkiloglu, byolacan}@havelsan.com.tr

Özet. Bir yazılım bileşeni ya da sisteminin yerine getirmesi gereken işlevler fonksiyonel gereksinimler ile tanımlanabilirken fonksiyonallık ile ilgili olmayan ve yazılımın nasıl olması gerektiğini belirleyen isterler, güvenilirlik, sürdürülebilirlik gibi kalite nitelikleri ve bu çalışmanın kapsamında olan kullanılabilirlik kapsamında belirlenmektedir. Kullanılabilirlik, ilgili yazılımın, son kullanıcı ile olan arayüzü vasıtasıyla kolay kullanılabilir olma yeteneği olarak tanımlandığında kullanılabilirlik testi söz konusu yeteneğin ne ölçüde açığa çıktığını belirleme amacı gütmektedir. Bu çalışmada kullanılabilirlik test yaklaşımının, belli kullanılabilirlik sezgisel yöntemleri kapsamında oluşturulmuş Yazılım Kullanılabilirlik Ölçüm Anketi, gözlemlenebilirlik ile beraber Sesli Düşün metodu kullanılarak bir kabin içi eğlence sistemi olarak geliştirilmekte olan SKYFE üzerinde uygulanması sonucu elde edilen deneyim paylaşılmaktadır.

Anahtar Kelimeler: Kullanılabilirlik, Yazılım Kullanılabilirlik Testi, Kabin İçi Eğlence Sistemi

Abstract. At one extreme, functional requirements define the functions that a software component or system should perform and on the other hand, non-functional requirements aimed at specifying how a software product is supposed to be are defined in terms of quality attributes such as reliability, maintainability and the one which comprises the scope of this study, namely, usability that is defined as the software capability of being easily used by the end user via the interface provided to them. Complementarily, usability testing targets at determining the extent to which this capability gets realized. In this study, the experience obtained by following the usability testing approach formed via Software Usability Measurement Questionnaire based on specific usability heuristics and Think Aloud method accompanied by observations is demonstrated in the scope of SKYFE In Flight Entertainment System that is in development.

Keywords: Usability, Software Usability Test, In Flight Entertainment System

1 Giriş

Son bir kaç on yılda geliştirilen ürünlerin kullanılabilirliğinin öneminin kavranmasıyla beraber (örneğin IBM ve Boing Co. gibi büyük ölçekli şirketlerin yazılım geliştirirken ve satın alırken kullanılabilirliği anahtar faktör olarak kabul etmeleri [1]) bir kalite faktörü olarak kullanılabilirliğin uluslararası standartlarda tanımlanması ihtiyacı ortaya çıkmıştır. Önceleri, bir kullanıcının bir sistemi ya da sistemin bir bileşenini ne kolaylıkla işletmeyi öğrenmesi, girdi sağlaması ve çıktıları yorumlaması olarak tarif edilirken [2] sonrasında, bir ürünün belli kullanıcılar tarafından belirli bir kullanım çerçevesinde belirlenmiş hedeflere etkililik, verimlilik ve memnuniyet kapsamında ulaşmak için ne ölçüde kullanılabildikleri şeklinde betimlenmiştir [3]. Yazılım bağlamında ise bir yazılım ürününün belli koşullar altında kullanıcı tarafından anlaşılma, öğrenilme, kullanılma ve ilgi çekme kabiliyeti olarak ifade edilmiştir [4]. Buradan hareketle belirlenecek gereksinimler baz alınarak ortaya konulacak ürünler kapsamında kullanıcıların üretkenliğinin artması, yazılım geliştiricilerin daha az efor harcamaları ve nihai ürünün destek ve bakım maliyetlerinin azalması beklenilebilecektir [5].

Kullanılabilirlik testi, köklerini klasik deneysel metodolojiden alan bir araştırma aracı olarak görülebilir [6]. Bu bağlamda, kullanılabilirlik testi, bir ürünün gerçek kullanıcıları tarafından denenmesi esnasında, kullanıcıların gözlemlenmesinin ve ürünün kullanımının onlar için kolay mı yoksa zor mu olduğu bilgisinin toplanmasının sistematik bir yolu olarak tanımlanabilir [7]. Yapılan bu tanımlarda, kullanılabilirliğin nasıl ölçüldüğüne dair bir açıklama görmüyoruz. Ancak ISO'nun "Kullanılabilirlik" [4] tanımı kullanılarak ISTQB tarafından şöyle bir tanımlama yapılmıştır: "Belirlenmiş koşullar altında yazılım ürününün kullanıcıya cazip geldiğini, kolay kullanılabildiğini, kolay öğrenilebildiğini ve anlaşılabilirliğini doğrulamak için yapılan test." [8]. Burada kullanılabilirliğin caziplik, kullanımda kolaylık, öğrenilebilirlik ve anlaşılabilirlik boyutları tanıma dahil edilmiştir. Bunların yanı sıra hatırlanabilirlik, hata ile karşılaşma sıklığı, etkinlik, etkililik ve tatmin boyutları da araştırmacılar tarafından problem uzayını tanımlamakta kullanılmışlardır [4][8][9].

Kullanılabilirlik boyutları, bir denek grubunun kullanımıyla ölçülebildiği gibi, bu işi kullanıcı bağımsız modeller ile değerlendirebilen yöntemler de mevcuttur. Bilişsel Modelleme yöntemleri olarak da bilinen bu yöntemlerin ilki GOMS (nam-ı diğer CMN GOMS) yöntemidir. Yöntem, araştırmacıları olan Card, Moran ve Newell'in [10] kullanıcı etkileşiminin mühendislik modelini belirleme çalışması sonrası ortaya çıkmıştır. Diğer bir yöntem olan Tuşlama Seviye Modeli (KLM) [11] ise tuşlama ve fare kullanımı ve kullanıcı düşünme sürelerini de dikkate alarak bir yazılım tasarımının kullanılabilirliğini değerlendirmektedir. Fitts Kanunu [12] adı verilen yöntem de bu modelleme çalışmalarına örnek olarak verilebilir. Bu yöntemde seçme görevlerinin hareket zamanlaması modellenerek kullanılabilirlik analizi gerçekleştirilmektedir.

Bu yöntemlerin yanı sıra, belirlenen sezgisel özellikler kullanılarak uygulamaya özel olarak oluşturulan senaryolar ve arkasından uygulanan anketler vasıtasıyla, ilgili uygulama hakkında kullanıcıların görüşlerini kullanan yöntemler de mevcuttur [13]. Bu çalışmada, yakın zamanda HAVELSAN A.Ş. ve THY Teknik A.Ş. tarafından ortaklaşa geliştirilmeye başlanan, yolcuları geleneksel koltuk arkası ekranlara

bağımlılıktan kurtararak kablosuz ağ yetenekli ve farklı işletim sistemleri üzerinde çalışabilen kişisel elektronik cihazlarına (örneğin tablet, dizüstü bilgisayar, akıllı telefon) İsteğe Bağlı Video ve Ses (Audio and Video On Demand), çevrim dışı internet ve elektronik yayın gibi hizmetler sunması beklenen [14] ilk yerli kablosuz Kabin İçi Eğlence Sistemi olan SKYFE'nin geliştirme sürecinde Nielsen'e ait sezgisel özniteliklerden ve kullanıcı görüşlerinden yararlanılarak yapılmış kullanılabilirlik analizi deneyimi raporlanmıştır [13].

Bu bildirin takip eden ikinci bölümünde uygulanan test metodolojisine değinilmektedir. Üçüncü bölüm, elde edilen test sonuçları ve bu sonuçların analizi üzerinedir. Son olarak, dördüncü bölüm bu deneyim sonrası SKYFE projesi kapsamında ulaşılan edinimleri özetlemektedir.

2 Test Yöntemi

SKYFE kullanıcı ara yüzünde yapılan değişikliklerin birbirini tekrar etmesinin gözlenmesini müteakip kullanılabilirlik ölçümünün SKYFE projesi kapsamında yapılması gerekliliği ortaya çıkmıştır. Yapılan araştırma sonucunda, son kullanıcıların değerlendirmeye dâhil edildiği Sezgisel Değerlendirme (Heuristic Evaluation) tekniğinin yaygın olarak kullanılması ve projede uygulanabilmesi için gerekli eforun olanaklı olması nedeniyle daha etkili ve kolay hayata geçirilebilir bulunmuştur.

SKYFE projesi bağlamında Nielsen'in sezgisel değerlendirme metodu kullanılmıştır. Bu metod kapsamında, ilk olarak SKYFE sistem özelliklerinin son kullanıcılar tarafından kullanılabilmesi için bir test senaryosu hazırlanmıştır. Test sırasında gözlemci teste müdahil olmadan test görev performanslarını ya da kullanıcı memnuniyetini önemli şekilde etkileyen vakaları kayıt altına alırken, kontrolör ise teste başlamadan kullanıcıları testin içeriği hakkında bilgilendirme, kullanıcılara test görevlerini sağlama, kullanıcıların sesli düşüncelerini yönetme, kullanıcıların test esnasında ne yapmayı amaçladıklarını belirtmeleri hususunda yönlendirme ve test oturumunun idame ettirilmesi sorumluluğunu icra etmiştir.

Hazırlanan test için çeşitli yaş, meslek ve mobil uygulama kullanma alışkanlıkları olan 16 kişilik bir grup oluşturulmuştur. Seçilen kullanıcılar, uygulamanın geliştirildiği platformlar için (IOS ve Android) eşit iki gruba ayrılmıştır. Katılımcılardan testi kendi başlarına ve uygulama hakkında bilgilendirme olmadan koşmaları istenmiştir. Test sırasında kullanıcıların geri dönüşleri gözlemlenmiş ve kaydedilmiştir.

Testin geçerli sonuçlar vermesini tehdit eden faktör kullanıcıların sesli düşünmekten çekinmesi durumudur. Bu faktörün etkisini en aza indirmek için kullanıcıların sistem hakkında düşüncelerini hiç çekincesiz ifade etmeleri yönünde gerekli telkin mekanizması işletilmiştir. Bunun yanı sıra, her bir platform için beş kişiden oluşan ilk kullanıcı kümesi ile elde edilen test sonuçları normal dağılım göstermemesinden dolayı, kullanıcı kümesinin genişletilerek her bir platform sekiz kişilik kullanıcı kümesiyle test edilmiştir.

Uygulanan test sonrasında kullanıcılara daha önceden Nielsen'in sezgisel değerlendirme metodlarına uygun olarak hazırlanmış olan anket uygulanmıştır. Anket soruları; sistem durumu görünürlüğü, sistem ve gerçek dünya arasındaki uyum,

kullanıcı kontrol ve özgürlüğü, tutarlılık ve standartlar, hatırlama yerine tanıma, esneklik ve kullanım verimliliği, estetik ve minimalist tasarım ve kullanıcılara tanıma, teşhis ve hatadan kurtulma yardımı sağlama başlıkları altında gruplanmıştır. Ankette bulunan sorular için, kullanıcılardan dereceleri "Hiç, %10, %25, %50, %85 ve Her Zaman" olan seçeneklerden sadece birini işaretlemeleri istenmiştir. Sezgisel değerlendirme grupları ve her bir grubun kapsamı ile en çok örtüşen anket soruları Tablo-1'de bulunmaktadır.

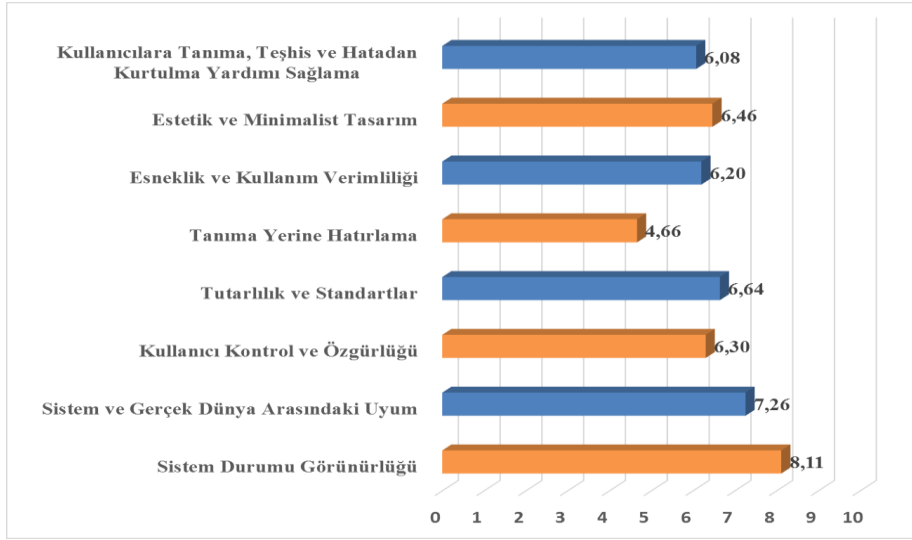
Tablo 1. Anket Soruları ve Sezgisel Öznitelik Grup Bilgileri

Soru No	Sezgisel Değerlendirme Grubu
Sistem Durumu Görünürlüğü (<i>Visibility of System Status</i>)	
1	Uygulamada neler olup bittiğini anlıyorum.
Sistem ve Gerçek Dünya Arasındaki Uyum (<i>Match Between System and The Real World</i>)	
2	Kullanılan ikonlara aşınayım. İkonlar kültürel değil, uluslararası yolcular için evrensel olarak tasarlanmış.
5	Menü içeriğe göre mantıklı kategorize edilmiş. Menü seçenekleri en mantıklı şekilde sıralanmış.
15	Bir dahaki uçuşumda, SKYFE'yi hatırlarım ve daha kolay kullanırım.
18	Müzik listemi yaratırken ve düzenlerken hiçbir sorunla karşılaşmadım.
Kullanıcı Kontrol ve Özgürlüğü (<i>User Control and Freedom</i>)	
4	Sayfalar arası geçişler kolay.
8	Kısa yol menü kullanışlı ve ana sayfada yönlendirmeler yeterli.
16	İstemediğim bir durumdan çıkmam ve geri dönmem mümkün.
Tutarlılık ve Standartlar (<i>Consistency and Standards</i>)	
9	Uygulamanın bütün ekranlarda formatı standart ve tutarlı.
10	SKYFE'nin renk kullanımı uygun ve başarılı.
11	SKYFE'nin yönlendirme dili tutarlı ve alışılabilir.
21	Benzer işlemleri benzer görevler için benzer elemanlar ile yapılır.
Hatırlama Yerine Tanıma (<i>Recognition Rather than Recall</i>)	
3	Seçilen ikonun, diğer ikonlara göre fark edilmesi kolay.
6	Buton başlıkları ve etiketler açık ve anlaşılır.
Esneklik ve Kullanım Verimliliği (<i>Flexibility and Efficiency of Use</i>)	
20	Bir filmin altyazısını ve dublajını kolayca ve hızlıca değiştirebiliyorum.
Estetik ve Minimalist Tasarım (<i>Aesthetic and Minimalist Design</i>)	
13	Her bir ekranda ihtiyacım olan bütün bilgiler mevcut.
14	Uçuşlarla ilgili verilen bilgiler kesin, tam ve anlaşılması kolay.
17	Görsellerin hizalanışı ve yerleri iyi dengelenmiş ve seçilmiş.
19	Müzik listemi düzenlediğim ekranda tasarım eksikliği ve gereksiz özellik yok.
Kullanıcılara Tanıma, Teşhis ve Hatadan Kurtulma Yardımı Sağlama (<i>Help Users Recognize, Diagnose, and Recover From Errors</i>)	
7	Uygulamada bir problem oluştuğunda anlaşılabilir. SKYFE bu problemleri önlemede yeterli ve yardımcı.
12	Hata mesajları bilgilendirme konusunda kısa ve öz.

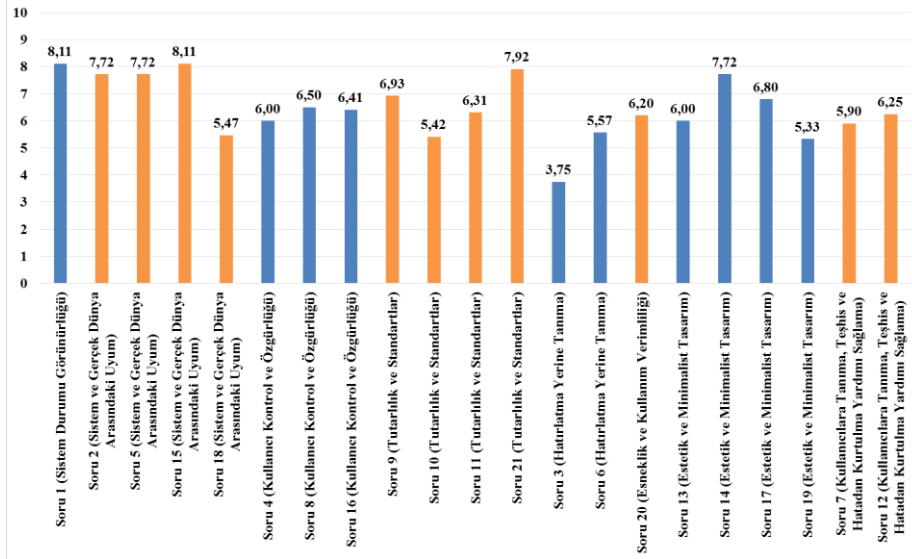
Anket sorularının sezgisel öznitelik odaklı sonuçlar vermesi için kullanıcının anlamakta güçlük çektiği sorular, ilgili özniteliklere göre anket esnasında açıklanmıştır.

3 Test Sonuçları

Her kullanıcının, test sonrası uygulanmak üzere hazırlanan ve içeriği Tablo-1’de verilen anket sorularını yanıtlaması sonucunda elde edilen veriler 10 puan üzerinden değerlendirilmiştir ve sorulara "Her Zaman" yanıtını veren kullanıcıların cevapları, kullanılabilirlik hatalarına odaklanma ve veri setindeki maksimum örnek değerleri olma noktasında hesaplamaya dâhil edilmemiştir. Anket sonuçları sezgisel öznelik bazlı olarak Şekil-1’de ve soru yönelimli olarak Şekil-2’de verilmektedir.



Şekil 1: Sezgisel öznelik bazlı değerlendirme ortalama sonuçları



Şekil 2. Soru yönelimli değerlendirme ortalama sonuçları

Anket sonuçları ve sesli düşün kullanılarak yapılan gözlemlerin ışığında, esneklik ve kullanım verimliliği niteliğinin görsellerin hizalanışı ve yerlerinin kullanıcılar tarafından beklendiği gibi olmamasından ötürü eksiklik gösterdiği fark edildi. Ek olarak bazı işlevlerin kullanıcı tarafından geç fark edildiği anlaşıldı. Bu nedenle görsellerin hizalanışına, yerlerinin tekrar ayarlanmasına ve geç fark edilen özelliklerin kullanıcıların alışık olduğu şekilde yönlendirme eklenmesi gerektiği anlaşıldı.

Kullanıcılara sunulan içeriğin bütün sayfalarda yeterli olduğu değerlendirilmesine karşın estetik ve minimalist tasarım niteliğini güçlendirmek için birkaç buton metninin sadeleştirilmesi gerektiği ortaya çıktı.

Kullanıcıların istemedikleri durumdan kolayca çıkabildiği gözlemlendi fakat bazı butonlara bastıklarını fark edememeleri ve menüye bakarak hangi içeriği görüntülediklerini anlamamaları sayfalar arası geçişi zorlaştırdığından uygulamada kullanıcı kontrol ve özgürlüğünü kısıtladığı ve kullanıcıların bazı eylemleri algılayamayarak kullanıcıların hatırlama yükünü artırabileceği ortaya çıktı. Basılan butonun ve seçilen menü öğesinin seçilmeyenlere göre ayırt edici olması için butonlara efektler eklenerek geri besleme sağlanması gerektiği anlaşıldı. Bazı buton etiketlerinin ise kullanıcılar tarafından kolay anlaşılması da kullanıcıların hatırlama yükünü artıracığı sonucu çıkarıldığından anlaşılmayan buton etiketlerinin değiştirilmesi gerektiği ortaya çıktı.

Tasarlanan ikonların kullanıcıların büyük bir çoğunluğu tarafından anlaşılması sonucu, sistem ve gerçek dünya arasında bu bağlamda uyumun sağlandığı ve ikonların evrensel olduğu değerlendirildi. Sadece bir ikonun kullanıcılar tarafından anlaşılması sesli düşün metoduyla ortaya çıktığından, o ikonun tekrar tasarlanmasına karar verildi. Kullanıcıların bir işlevi yaptıktan sonra benzer işlevleri daha kolay yapabilmeleri uygulamada benzer işlevlerin benzer elemanları ve yöntemleri kullandığını ve yönlendirme dilinin tutarlı olduğunu fark ettirdi.

Hata mesajlarından birkaçının kullanıcıları yanlış yönlendirebildiğinin gözlenmesi kullanıcılara hatadan kurtulma yardımı sağlama açısından eksiklik olduğunu gösterdiğinden bu hata mesajlarının metinlerinin daha öz ve anlaşılır olma gereği ile beraber yönlendirmelerin sadeleştirilmesi ihtiyacı tespit edildi.

4 Sonuç

Bu çalışmada, HAVELSAN A.Ş. ve THY Teknik A.Ş. tarafından ortaklaşa geliştirilmeye başlanan ilk yerli kablosuz Kabin İçi Eğlence Sistemi olan SKYFE'nin, belli kullanıcılar tarafından, belirli bir kullanım çerçevesinde, belirlenmiş olan hedeflere ulaşmak için etkililik, verimlilik ve memnuniyet kapsamında kullanılabilirliği araştırılmıştır.

Araştırma sonunda kullanıcıların yazılı ve sözlü ifadelerinden, olumlu ve olumsuz gözlemlerinin yanı sıra beklentileri de saptanabilmektedir. Araştırma sonucu elde edilen bulgular doğrultusunda SKYFE'nin içerik ve tasarımı ile ilgili düzenlemelerin yapılabilmesi ve kullanılabilirliğinin artırılabilmesi için değişiklik önerilerinde bulunulmuştur. Kullanılabilirlik testinden elde edilen bulgular ve öneriler

kapsamında, SKYFE'nin grafiksel kullanıcı arayüz tasarımının etkilenen kısımları için uyumsuzluklar açılmış ve bunların çözümüyle beraber kullanılabilirlik kapsamında SKYFE yazılımının olgunlaştırılması sağlanmıştır. Buradan hareketle benzer yazılım projelerinde de bu araştırmadan elde edilen bulgulardan ve önerilerden yararlanılmasının faydalı olacağı öngörülmüştür.

Kaynaklar

1. Capilla, Rafael, Laura Carvajal, and Hui Lin. "Addressing Usability Requirements in Mobile Software Development." *Relating System Quality and Software Architecture* (2014): 303-324.
2. IEEE (1990). *Standard Glossary of Software Engineering Terminology*. IEEE Standard 610.12-1990.
3. ISO, SFSEN. "9241-11. 1998." *Ergonomic Requirements for Office Work with Visual Display Terminals (VDTs)–Part II Guidance on Usability*.
4. ISO/IEC 9126-1 (2001). *Software Engineering – Product Quality – Part 1: Quality Model*. International Organization for Standardization.
5. Juristo, N., Helmut, W., Windl, H., & Constantine, L., "Introducing Usability", *IEEE software* 18.1 (2001): 20-21.
6. Rubin, J. (1994). *Handbook of usability testing: How to plan, design, and conduct effective tests*. New York, NY: John Wiley & Sons, Inc.
7. Dumas, J.S. & Redish, J.C. (1993). *A practical guide to usability testing*. Norwood, NJ: Ablex Publishing Corporation.
8. Yazılım Test ve Kalite Derneği Terimler Sözlüğü Çalışma Grubu, "ISTQB Yazılım Testi Terimler Sözlüğü Versiyon 1.0", Turkish Testing Board (2014).
9. Nielsen, J. (1994). *Usability Engineering*, Academic Press Inc, p 165
10. Stuart K. Card , Allen Newell , Thomas P. Moran, *The Psychology of Human-Computer Interaction*, L. Erlbaum Associates Inc., Hillsdale, NJ, 1983
11. Card, Stuart K., Thomas P. Moran, and Allen Newell. "The keystroke-level model for user performance time with interactive systems." *Communications of the ACM* 23.7 (1980): 396-410.
12. Fitts, P. M., & Radford, B. K. (1966). Information capacity of discrete motor responses under different cognitive sets. *Journal of Experimental Psychology*, 71, 475-482.
13. Nielsen, Jakob. "10 Usability Heuristics for User Interface Design." Fremont: Nielsen Norman Group.[Consult. 20 maio 2014]. Disponível na Internet (1995).
14. HAVELSAN A.Ş. Basın Duyuruları, <http://www.havelsan.com.tr/Basin/BasinDuyurulari.aspx?id=283>, 18.11.2014