

PERBANDINGAN METODE EVALUASI *USABILITY* (STUDI KASUS : PENGGUNAAN PERANGKAT *SMARTPHONE*)

Andrie Pasca Hendradewa

Jurusan Teknik Industri, Fakultas Teknologi Industri, Universitas Islam Indonesia
Jl. Kaliurang Km 14.5, Ngemplak, Sleman, Yogyakarta, 55584
Email : andrie.pasca@uii.ac.id

ABSTRACT

Usability is a key factor that determine success of a product or interactive system, like smartphone. The increasing use of smartphone and its rapid technology development, requires usability evaluation method that is more accurate and effective to found usability problem, so it can be used for product improvement in development process. This study suggests a usability evaluation method that is most effective to identify problem in smartphone usage. Three usability evaluation methods that commonly used are compared in this study, including: think-aloud evaluation (TA), cognitive walkthrough (CW), and heuristic evaluation (HE). This study involved 15 smartphone users as participants and 16 usability experts who is also a smartphone users as evaluators. The results showed, among the three usability evaluation methods, heuristic evaluation (HE) has the highest effectiveness and severity methods, while cognitive walkthrough (CW) and think-aloud evaluation (TA) still has potential for further development.

Keywords : Smartphone, Usability, User Interface Design, Metode Evaluasi Usability.

1. PENDAHULUAN

Smartphone merupakan perangkat telekomunikasi yang semakin umum digunakan dalam aktivitas keseharian saat ini. Pada kuartal ke-4 tahun 2014, berdasarkan data yang diperoleh dari nielsen.com, sebanyak 76.3% pria dan 77.8% wanita di Amerika merupakan pengguna *smartphone* (*The Nielsen Company*, 2014). Data statistik dari *businessinsider.com* juga menunjukkan bahwa jumlah pengguna *smartphone* di berbagai negara diprediksi mengalami peningkatan pada tahun 2015 (*Businessinsider*, 2014). Hal yang paling mencolok dari karakteristik *smartphone* adalah penggunaan teknologi layar sentuh sebagai media input. Penggunaan tangan sebagai sarana input merupakan metode yang atraktif untuk menghasilkan sebuah interaksi manusia - komputer yang alamiah (Wang & Ren, 2009).

Interaksi dari layar sentuh merupakan interaksi paling “langsung” antara manusia dengan komputer, dimana tampilan informasi dan kendali berada dalam satu permukaan (Albinsson & Zhai, 2003). Di

dalam penelitian tersebut juga dijelaskan bahwa layar sentuh juga memiliki beberapa keterbatasan antara lain : (1) jari, tangan, dan lengan dapat menjadi penghalang sebagian dari layar, (2) jari manusia sebagai alat penunjuk memiliki “resolusi” yang rendah sehingga menyulitkan untuk menunjuk target yang lebih kecil dari lebar jari. Keterbatasan-keterbatasan layar sentuh sebagai media input tentu berpotensi menimbulkan kendala dalam menggunakan perangkat *smartphone*. Oleh karena itu, tingkat kemudahan dalam menggunakan perangkat tersebut merupakan salah satu aspek penting yang perlu diperhatikan dalam perancangan produk *smartphone*.

Usability saat ini dikenal sebagai kunci utama yang menentukan keberhasilan dari sebuah sistem interaktif atau produk (Maguire, 2001). *Usability* disebut juga sebagai karakteristik dari kualitas produk dimana subkarakteristiknya terdiri dari kemampuan untuk dapat dikenali secara sesuai (*appropriate recognizability*), kemudahan dipelajari (*learnability*), kemudahan dioperasikan (*operability*), perlindungan kesalahan oleh pengguna (*user*

error protection), keindahan tampilan antarmuka (*user interface aesthetics*), dan aksesibilitas (*accessibility*) (ISO / IEC 25010, 2011). Sistem yang dirancang dengan tidak baik dan tidak mudah digunakan akan membuat pengguna kesulitan mempelajari dan mengoperasikan, yang berimbas pada sistem tersebut jarang atau salah digunakan sehingga biaya yang harus ditanggung oleh organisasi pengguna sistem menjadi tinggi dan berbahaya bagi reputasi perusahaan yang mengembangkan sistem tersebut (Maguire, 2001). Pentingnya *usability* menunjukkan bahwa *usability evaluation* merupakan proses penting dalam tahap pengembangan untuk menghasilkan sistem atau produk yang lebih baik dari sebelumnya. Metode evaluasi *usability* atau yang juga disebut sebagai *Usability Evaluation Method* (UEM) terbagi dua metode antara lain inspeksi berbasis ahli atau *expert - based inspection* dan pengujian berbasis pengguna atau *user-based testing*. Metode inspeksi melibatkan ahli *usability* untuk menemukan masalah *usability* dengan mengikuti petunjuk, heuristik, atau penelusuran. Sedangkan melalui metode pengujian, permasalahan *usability* ditemukan dengan cara melakukan observasi kepada pengguna pada saat menggunakan produk atau sistem (Nielsen & Mack, 1994).

Namun permasalahan saat ini adalah, metode evaluasi *usability* saat ini tidak mempertimbangkan karakteristik dari perangkat *mobile* berbasis layar sentuh dan antarmuka fisik. Sebagai salah satu solusi atas permasalahan tersebut, telah terdapat usulan 11 heuristik yang merupakan penyesuaian prinsip *usability* untuk metode *heuristic evaluation* dalam konteks evaluasi *touchscreen-based mobile device* (Inostroza dkk, 2012). Namun belum diketahui apakah metode tersebut sudah lebih baik daripada metode-metode tradisional lainnya (eg: *cognitive walkthrough* and *think-aloud evaluation*), sehingga perlu ada pembuktian dengan membandingkan metode tersebut terhadap metode lainnya. Penelitian ini membandingkan tiga buah metode evaluasi *usability* (UEM) untuk dapat memberikan rekomendasi metode evaluasi yang paling

efektif digunakan untuk menemukan permasalahan kemudahan perangkat *smartphone* beserta jumlah evaluator atau partisipan yang diperlukan.

2. METODE PENELITIAN

Penelitian ini membandingkan tiga metode yang paling banyak digunakan dalam tahap pengembangan untuk menemukan permasalahan kemudahan, yaitu : *Think - Aloud Evaluation* (TA), *Cognitive Walkthrough* (CW), dan *Heuristic Evaluation* (HE). Ketiga metode tersebut, dalam penelitian ini, digunakan untuk menemukan permasalahan kemudahan pada perangkat *smartphone*, sedangkan metode evaluasi lain yang bersifat mengukur tingkat kemudahan tidak diikuti sertakan.

2.1. *Think - Aloud Evaluation* (TA)

TA merupakan metode pengujian berbasis pengguna yang melibatkan *end user* untuk melakukan verbalisasi secara kontinyu terhadap apa yang dipikirkan saat menggunakan sistem. Dengan melakukan verbalisasi, memungkinkan pengamat untuk menginterpretasikan pada bagian antarmuka mana yang memiliki masalah (Nielsen, 1993). Pada saat pengguna melakukan verbalisasi, seluruh komentar direkam, sehingga semua yang dipikirkan oleh pengguna dapat ditangkap dan poin - poin penting tidak terlewat pada saat proses analisis (Galitz, 2002). Keluaran dari metode ini berupa protokol verbal dan masalah *usability* pada tahap tertentu disiklus interaksi pengguna (Jaspers, 2009).

2.2. *Cognitive Walkthrough* (CW)

CW merupakan metode inspeksi yang berfokus pada kemudahan desain untuk dipelajari melalui penelusuran (Wharton dkk, 1994). Metode ini dilakukan oleh *expert evaluator* dengan membutuhkan tugas - tugas yang representatif, daftar lengkap urutan penyelesaian tugas, dan gambaran asumsi latar belakang calon pengguna. Dalam melakukan *walkthrough* pada setiap *action sequence*, *expert evaluator* perlu

menjawab empat buah pertanyaan sebagai berikut (Wharton dkk, 1994) :

- Akankah pengguna mencoba untuk mendapatkan efek yang benar?
- Akankah pengguna mengetahui tersedianya tindakan yang tepat?
- Akankah pengguna mengasosiasikan tindakan yang tepat terhadap efek yang ingin didapatkan?
- Jika tindakan yang benar dilakukan, akankah pengguna melihat adanya progress untuk mencapai solusi?

Keluaran dari metode CW adalah berupa list masalah dan potensi masalah *usability* pada tahap tertentu di dalam siklus interaksi pengguna (Jaspers, 2009).

2.3. Heuristic Evaluation (HE)

HE merupakan metode inspeksi yang melibatkan ahli *usability* untuk menilai apakah elemen - elemen yang ada di suatu sistem mengikuti prinsip - prinsip *usability* (Nielsen & Mack, 1994). Metode ini dilakukan oleh masing - masing *evaluator* yang melakukan inspeksi terhadap suatu antarmuka secara individu. Komunikasi antar *evaluator* hanya diperbolehkan setelah evaluasi selesai dilakukan. Peraturan ini diterapkan untuk memastikan hasil evaluasi adalah hasil yang *independent* dan tidak bias. Selama pelaksanaan evaluasi, *evaluator* akan memeriksa antarmuka dalam beberapa kali, melakukan inspeksi terhadap bermacam elemen interaktif, dan membandingkannya dengan prinsip *usability* yang ada (Holzinger, 2005).

Heuristik yang dikemukakan oleh Nielsen (Nielsen & Mack, 1994) merupakan salah satu prinsip *usability* yang sering digunakan untuk melakukan inspeksi antarmuka suatu produk. Dalam konteks *touchscreen - based mobile device*, dengan mengembangkan metode heuristik oleh Nielsen, Inostroza dkk (2012) mengusulkan 11 *heuristics* sebagai berikut :

- Keterlihatan status sistem.
- Kecocokan antara sistem dan realita.

- Kendali dan kebebasan pengguna.
- Konsistensi dan standar.
- Pencegahan kesalahan.
- Minimalisir beban memori pengguna.
- Kustomisasi dan jalan pintas.
- Estetika dan desain minimalis.
- Bantuan untuk mengenali, mendiagnosa, dan memperbaiki kesalahan.
- Bantuan dan dokumentasi.
- Interaksi fisik dan ergonomi.

Keluaran dari pelaksanaan metode HE adalah berupa daftar heuristik yang dilanggar beserta *severity rating* nya (Jaspers, 2009). Pelaksanaan evaluasi dibagi ke dalam tiga grup dengan menggunakan metode evaluasi yang berbeda. Grup pertama menggunakan TA method dengan 15 partisipan yang merupakan pengguna mahir perangkat *smartphone*. Seluruh partisipan tersebut merupakan mahasiswa tingkat master di salah satu perguruan tinggi. Dalam penelitian ini, kami mengasumsikan bahwa kemampuan pengoperasian perangkat dan kemampuan verbalisasi seluruh TA partisipan adalah sama. Grup kedua dan ketiga berturut-turut menggunakan CW metode dan HE metode dengan masing - masing grup terdiri dari 8 *evaluator* yang memiliki keahlian dibidang *user interface* dan juga mahir dalam menggunakan perangkat *smartphone*. Dalam penelitian ini, pengetahuan dan keahlian seluruh *evaluator* diasumsikan sama.

Permasalahan *usability* gabungan yang ditemukan dari ketiga grup kemudian diberikan penilaian *severity rating* oleh 3 orang ahli *user interface* atau yang selanjutnya disebut dengan *rater*. *Severity rating* yang digunakan dalam penelitian ini merupakan *skala rating* usulan Nielsen yang memiliki *skala rating* 0 sampai 4 (Nielsen, 2014). Penilaian *severity rating* diperlukan untuk menentukan bobot untuk setiap temuan masalah *usability*.

Untuk membandingkan ketiga metode evaluasi tersebut, penelitian ini menggunakan beberapa ukuran perbandingan antara lain *validity*, *thoroughness*, *effectivity*, dan *severity rating*

metode (Hartson dkk, 2003). Beberapa ukuran perbandingan metode evaluasi tersebut dijelaskan sebagai berikut:

- Validitas merupakan ukuran yang menentukan seberapa baik metode evaluasi memberikan hasil yang diinginkan, UEM dengan validitas rendah akan menemukan banyak masalah yang tidak relevan. Formulasi untuk menentukan nilai *validity* adalah sebagai berikut: (dimana *P* adalah jumlah aspek yang teridentifikasi sebagai masalah dan *A* adalah jumlah masalah nyata yang ada).

$$Validity = |P \cap A| / |P| \dots \dots \dots (1)$$

- *Thoroughness* atau ketelitian mengindikasikan proporsi masalah nyata atau *real problem* yang terdapat pada rancangan antarmuka. UEM dengan ketelitian rendah berarti tidak dapat mengidentifikasi masalah usability yang penting. *Thoroughness* atau ketelitian suatu UEM diformulasikan sebagai berikut :

$$Thoroughness = |P \cap A| / |A| \dots \dots \dots (2)$$

- Efektifitas didefinisikan sebagai gabungan dari validitas dan ketelitian. Efektifitas UEM memiliki jangkauan nilai dari 0 hingga 1. Jika nilai ketelitian dan validitas rendah, maka nilai efektifitas juga rendah.

$$Effectifity = Thoroughness \times Validity \dots \dots \dots (3)$$

- *Severity rating* metode merupakan hasil rata - rata dari nilai *severity rating* atau tingkat urgensi masalah yang ditemukan oleh metode evaluasi, yang akan memperlihatkan sejauh mana suatu UEM dapat menemukan masalah yang penting. Nilai tersebut dihitung dengan formulasi berikut: (dimana $\sum s(rpf)$

adalah total *severity rating* yang ditemukan oleh UEM)

$$Severity\ rating\ method = \sum s(rpf) / \text{jumlah masalah nyata yang ditemukan} \dots \dots \dots (4)$$

Penelitian dilakukan dengan meminta partisipan atau *evaluator* melakukan evaluasi *smartphone* dengan mengikuti beberapa tugas representatif yang ditentukan berdasarkan fitur - fitur umum yang ada pada *smartphone*. Sebagai batasan, tugas yang berhubungan dengan jenis aplikasi tertentu (*chat, music, social media, etc*) tidak diikuti sertakan. Kategori tugas - tugas representatif yang dipilih antara lain :

- Penambahan kontak : Partisipan diminta untuk menambahkan kontak, lengkap dengan informasi lainnya beserta penambahan grup kontak ke dalam perangkat.
- Pemanggilan kontak : Partisipan diminta melakukan panggilan telepon kepada kontak yang telah dimasukkan sebelumnya.
- Penerimaan panggilan: Partisipan diminta untuk menjawab panggilan telepon.
- Pengiriman pesan : Partisipan diminta untuk mengirimkan pesan singkat SMS sesuai dengan detail informasi yang diberikan.
- Pengambilan gambar : Partisipan diminta untuk mengambil gambar dengan menggunakan kamera *smartphone*.
- Sambungan mobile data : Partisipan diminta untuk melakukan pengaturan dan aktivasi koneksi internet melalui kartu SIM.
- *Hotspot tethering* : Partisipan diminta melakukan pengaturan dan aktivasi fitur *hotspot tethering*.
- Pengecekan ketersediaan memori: Partisipan diminta untuk melakukan pengecekan memori yang masih tersedia diperangkat.

Tabel 1. Jumlah temuan masalah per UEM

<i>Participant / Evaluator ID</i>	Jumlah Temuan Masalah		
	<i>TA</i>	<i>CW</i>	<i>HE</i>
1	8	19	28
2	8	9	15
3	8	6	4
4	3	4	5
5	3	3	2
6	3	2	1
7	4	2	0
8	4	1	1
9	1		
10	4		
11	2		
12	0		
13	0		
14	1		
15	0		
Total	49	46	56

- Instalasi aplikasi : Partisipan diminta untuk melakukan instalasi aplikasi melalui jaringan internet.
- Penghapusan aplikasi : Partisipan diminta untuk melakukan *uninstall* aplikasi yang terdapat pada perangkat.
- Sambungan *WiFi* : Partisipan diminta untuk melakukan pengaturan dan aktivasi koneksi *WiFi*.
- Pengunduhan *file* : Partisipan diminta untuk melakukan aktivitas pengunduhan *file* dari internet.
- Pengorganisasian *file* : Partisipan diminta untuk mencari *file*, memindahkan *file*, dan membuat *folder*.
- *Screenshot* : Partisipan diminta mengambil gambar tampilan halaman antarmuka dengan fitur *screenshot*.

Pada penelitian ini, *smartphone* yang digunakan memiliki sistem operasi Android. Android merupakan sistem operasi yang mendominasi pasar *smartphone* di seluruh dunia (IDC, 2015). Salah satu ciri khas perangkat *smartphone* berbasis Android adalah *user interface* yang mengalami customisasi untuk setiap merek (contoh :

Perbandingan *User Interface* Android (Androidpit, 2015)). Diantara sistem operasi Android yang ada, versi *platform* yang paling banyak digunakan adalah versi 4.4 (KitKat) (Statista, 2014). Berdasarkan data statistik tersebut, perangkat *smartphone* yang diperlukan sebagai obyek evaluasi dalam penelitian ini merupakan perangkat *smartphone* yang berbasis pada *platform* Android KitKat.

3. HASIL DAN PEMBAHASAN

3.1.1. Hasil

Evaluasi *usability* dilakukan dengan mengikuti protokol pelaksanaan untuk masing - masing metode. Temuan - temuan masalah dari seluruh partisipan maupun *evaluator* pada setiap grup kemudian diperiksa untuk mengetahui masalah-masalah baru yang ditemukan oleh setiap partisipan maupun *evaluator*. Tabel 1 menampilkan hasil evaluasi *smartphone* dengan menggunakan UEM yang berbeda dan menunjukkan bahwa jumlah temuan masalah *usability* berbeda - beda untuk setiap UEM. Diantara ketiga grup tersebut, *heuristic evaluation* merupakan metode yang

Tabel 2. Jumlah *Real Problem* per UEM

Metode	Problem	Jumlah Masalah Berdasarkan Severity Rating					Real Problem
		0	1	2	3	4	
TA	49	6	17	11	10	5	43
CW	46	4	16	10	11	5	42
HE	56	4	16	14	15	7	52

0: not a problem, 1: cosmetic problem, 2: minor problem, 3: major problem, 4: catastrophe

Tabel 3. Validitas Metode Evaluasi *Usability*

Metode	Problem	Real Problem	Validity (%)
TA	49	43	87.76
CW	46	42	91.3
HE	56	52	92.86

Tabel 4. Ketelitian Metode Evaluasi *Usability*

Metode	Real Problem	Real Problem Total	Thoroughnes(%)
TA	43	82	52.44
CW	42		51.22
HE	52		63.41

Tabel 5. Efektifitas Metode Evaluasi *Usability*

Metode	Validity	Thoroughness	Effectivity (%)
TA	0.8776	0.5244	46.02
CW	0.913	0.5122	46.76
HE	0.9286	0.6341	58.88

Tabel 6. *Severity Rating* per Metode Evaluasi

Metode	Total Severity Rating of Real Problem	Real Problem	Severity Rating Method
TA	90	43	2.093
CW	90	42	2.143
HE	122	52	2.346

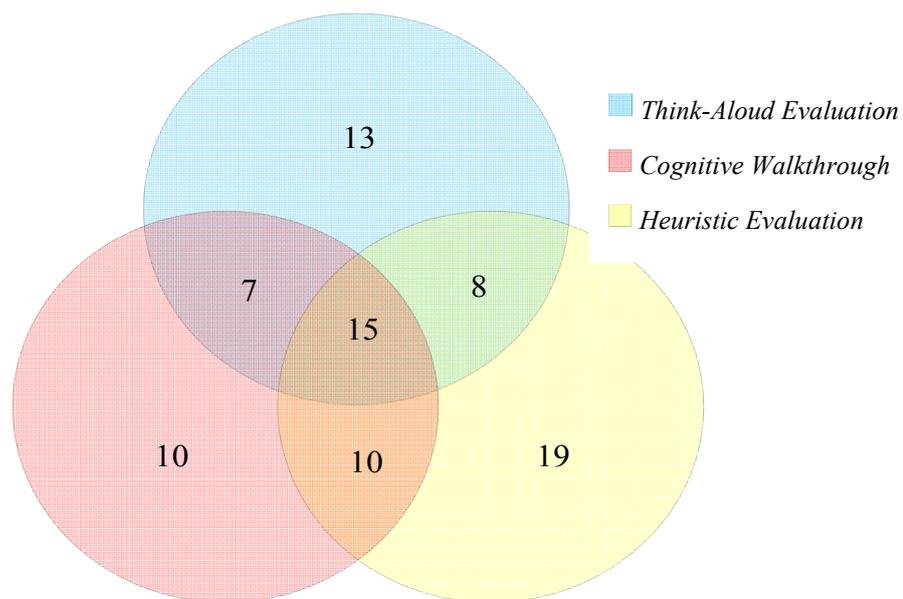
Tabel 7. Rangkuman Hasil Penilaian Metode Evaluasi (UEM)

Ukuran Perbandingan	UEM (Metode Evaluasi)		
	TA	CW	HE
Validity	87.76%	91.3%	92.86%
Thoroughness	52.44%	51.22%	63.41%
Effectivity	46.02%	46.67%	58.88%
Severity Rating Method	2.093	2.143	2.346

dapat menghasilkan temuan masalah *usability* paling banyak dibandingkan metode lainnya. Temuan-temuan masalah dari ketiga UEM tersebut kemudian digabungkan ke dalam satu daftar untuk dapat diberikan penilaian *severity rating*. Dalam proses penggabungan, 1 masalah *usability* yang juga ditemukan oleh lebih dari 1 UEM diperhitungkan sebagai 1 masalah yang sama. Sehingga dalam daftar masalah hasil penggabungan ketiga UEM yang akan di-rating, tidak terdapat masalah yang berulang.

Pada tabel 2 ditunjukkan hasil *severity rating* untuk setiap masalah *usability* yang ditemukan oleh ketiga UEM. *Severity rating* akhir merupakan hasil pembulatan dari rata-rata *rating* dari ketiga *rater*. Masalah yang memiliki nilai *severity rating* 1 sampai 4 dikategorikan sebagai *real problem*.

Gambar 1 menunjukkan jumlah temuan masalah gabungan dari ketiga UEM yaitu sebanyak 82 *real problem*. Irisan antar diagram menunjukkan bahwa terdapat sejumlah *real problem* yang juga ditemukan oleh lebih dari satu UEM.



Gambar 1. Diagram Venn untuk *Real Problem* dari Tiga Metode Evaluasi.

Berdasarkan data pada tabel 2 dan gambar 1 tersebut, jumlah *real problem* yang ditemukan oleh UEM dan jumlah *real problem* yang ada dalam sistem dapat digunakan untuk mengetahui nilai validitas, *thoroughness*, and *effectivity* masing-masing UEM seperti yang dapat dilihat pada tabel 3, 4, dan 5.

Hasil penilaian *severity rating* untuk setiap masalah *usability* yang ditemukan oleh ketiga UEM dapat digunakan untuk melihat nilai *severity rating method*. Tabel 6 menunjukkan nilai *severity rating method* untuk ketiga UEM.

Beberapa ukuran perbandingan telah didapatkan melalui pelaksanaan evaluasi

dengan UEM yang berbeda-beda, baik yang melibatkan partisipan maupun *evaluator*. Ukuran perbandingan metode evaluasi untuk ketiga UEM dirangkum dalam tabel 7.

Berdasarkan tabel 7, metode *heuristic evaluation* (HE) memiliki nilai *validity* dan *thoroughness* yang paling tinggi dibandingkan dengan *think-aloud evaluation* (TA) dan *cognitive walkthrough* (CW). Semakin tinggi nilai *validity* berarti semakin besar proporsi *real problem* jika dibandingkan dengan jumlah keseluruhan temuan masalah oleh suatu UEM. Tingginya nilai *validity* dan *thoroughness* metode *heuristic evaluation* (HE) membuat *effectivity* metode tersebut bernilai besar.

Dengan kata lain, dalam konteks evaluasi *usability smartphone*, UEM yang paling efektif adalah *heuristic evaluation*.

Severity rating method menunjukkan seberapa penting masalah *usability* yang ditemukan oleh suatu UEM. Semakin besar nilai *severity rating method*, maka semakin baik suatu UEM dalam mengidentifikasi masalah - masalah dengan tingkat urgensi tinggi. Berdasarkan tabel 7, metode *heuristic evaluation* memiliki *severity rating* metode yang paling tinggi dibandingkan dengan metode *think - aloud evaluation* dan *cognitive walkthrough*. Namun, *severity rating* metode untuk ketiga UEM jika dibulatkan, semuanya menghasilkan nilai = 2 (nilai *severity rating* = 2 merupakan minor *usability problem* : prioritas rendah untuk dilakukan perbaikan (Nielsen, 2014)).

3.1.2. Pembahasan

Penelitian ini menunjukkan bahwa dari perbandingan ketiga metode evaluasi *usability*, metode *heuristic evaluation* memiliki nilai yang paling baik hampir di seluruh ukuran perbandingan. Selain itu, penelitian ini juga memberikan validasi terhadap 11 heuristik usulan Inostroza dkk (2012), bahwa dengan mengikut sertakan aspek *physical interaction and ergonomics*, dalam konteks evaluasi teknologi *touchscreen - based mobile device*, metode *heuristic evaluation* dapat menghasilkan temuan masalah yang lebih banyak dan memiliki nilai efektifitas yang jauh lebih tinggi dibandingkan metode *cognitive walkthrough* dan *think - aloud evaluation*. Selain itu, juga diketahui bahwa metode *cognitive walkthrough* dan *think - aloud evaluation* memiliki nilai yang “kurang baik” hampir diseluruh ukuran apabila dibandingkan dengan metode *heuristic evaluation*. Hal tersebut dikarenakan sampai saat ini kedua metode tersebut belum terdapat pengembangan atau penyesuaian secara spesifik untuk konteks evaluasi teknologi *smartphone*.

Beberapa hasil mengenai perbandingan metode evaluasi *usability* untuk perangkat *smartphone* telah ditemukan, namun penelitian ini memiliki keterbatasan yaitu partisipan dan *evaluator* yang dilibatkan seluruhnya berasal dari Indonesia, dimana kebanyakan masyarakat belum terbiasa untuk melakukan *think - out - loud*. Oleh karena itu, hasil evaluasi *usability* mungkin dapat berbeda apabila partisipan dan *evaluator* yang dilibatkan berasal dari negara yang cenderung familiar dengan budaya *think - out - loud*.

4. KESIMPULAN

Penelitian ini mengusulkan bahwa UEM yang paling disarankan untuk digunakan dalam proses evaluasi *usability* perangkat *smartphone* adalah *heuristic evaluation*. Secara lebih jauh, heuristik yang digunakan adalah heuristik usulan Inostroza dkk (2012) yang telah mengalami penyesuaian untuk konteks teknologi *touchscreen - based mobile device*. Namun penelitian lebih jauh diperlukan untuk mengembangkan penelitian ini. Penelitian selanjutnya dapat menemukan nilai perkiraan yang sesuai untuk menemukan estimasi jumlah partisipan atau *evaluator* yang dibutuhkan oleh UEM dalam konteks evaluasi *usability smartphone*. Berdasarkan penelitian ini juga dapat dilihat bahwa masih terdapat peluang dikembangkannya metode *cognitive walkthrough* dan *think - aloud evaluation* agar lebih efektif digunakan untuk evaluasi *usability* dalam konteks teknologi *smartphone*. Oleh karena itu, pengembangan metode evaluasi *usability*, khususnya metode *cognitive walkthrough* dan *think - aloud evaluation* dapat dilakukan pada penelitian selanjutnya.

DAFTAR PUSTAKA

- Albinsson, P.A and Zhai, S. 2003. High Precision Touch Screen Interaction. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, p. 105-112.
- Androidpit. 2015. Android UI Comparison (<http://www.androidpit.com/android-ui-comparison> diakses pada 10 July 2015).
- Businessinsider. 2014. Smartphone Adoption Platform And Vendor Trends In Major Mobile Markets Around World 2015 <http://www.businessinsider.co.id/smartphone-adoption-platform-and-vendor-trends-in-major-mobile-markets-around-world-2015-3> diakses pada 12 November 2014).
- Galitz, W.O. 2002. The Essential Guide to User Interface Design: An Introduction to GUI Design Principles and Techniques (Second Edition). John Wiley & Sons Inc.
- Hartson, H.R, Andre, T.S, Williges, R.C. 2003. *Criteria for Evaluating Usability Evaluation Methods*. International Journal of Human-Computer Interaction, vol (13) 4, p. 145-181.
- Holzinger, A. 2005. Usability Engineering Methods for Software Developers. Communication of The ACM, vol 11 (1).
- IDC International Data Corporation. 2015. Smartphone OS Market Share (<http://www.idc.com/prodserv/smartphone-os-market-share.jsp> diakses pada 10 July 2015).
- Inostroza, R., Rusu, C., Roncagliolo, S., Jimenez, C., Rusu, V. 2012. *Usability Heuristics for Touchscreen - Based Mobile Devices*. 2012 9th International Conference of Information Technology New Generation, p. 662-667.
- ISO / IEC 25010. 2011. Systems and Software Quality Requirements and Evaluation (SQuaRE) – System and Software Quality Models. International Organization for Standardization, Geneva, Switzerland.
- Jaspers, M.W.M. 2009. *A Comparison of Usability Methods for Testing Interactive Health Technologies: Methodological Aspects and Empirical Evidence*,” International Journal of Medical Informatics, vol 78, p. 340-353.
- Maguire, M. 2001. *Method to Support Human-Centered Design*. International Journal of Human-Computer Studies, vol. 55 (4), p. 587-634.
- Nielsen, J. 1992. *Finding Usability Problems Through Heuristic Evaluation*. Proceedings of the SIGCHI conference on human factors in computing systems, p. 373-380.
- Nielsen, J. 2013. Usability Engineering. Boston: Academic Press Inc.
- Nielsen, J. 2014. Severity Ratings for Usability Problems (<http://www.nngroup.com/articles/how-to-rate-the-severity-of-usability-problems/> diakses pada 15 Nov. 2014).
- Nielsen, J., Landauer, T.K. 1993. *A Mathematical Model of The Finding of Usability Problems*. Proceedings of ACM Inter CHI'93 Conference, p. 206-213.
- Nielsen, J., Mack, R.L. 1994. Usability Inspection Methods, New York: John Wiley & Sons, USA.
- Statista, 2014. Share of Android Platforms on Mobile Devices with Android OS (<http://www.statista.com/statistics/271774/share-of-android-platforms-on-mobile-devices-with-android-os/> diakses pada 21st Nov.

- The Nielsen Company. 2014. Smartphone Owners Are As Diverse As Their Devices
(<http://www.nielsen.com/us/en/insights/news/2015/smartphone-owners-are-as-diverse-as-their-devices.html>) diakses pada 21 Nov.
- Wang F. and Ren. X. 2009 *Empirical Evaluation for Finger Input Properties in Multi-Touch Interaction*. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, p. 1063-1072.
- Wharton, C., Rieman, J., Lewis, C., Polson, P. 1994. The Cognitive Walkthrough Method: A Practitioner's Guide," Institute of Cognitive Science University of Colorado: Colorado.